DEVELOPMENT OF AN AMINE DEHYDROGENASE

A Dissertation
Presented to
The Academic Faculty

by

Michael J. Abrahamson

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemical and Biomolecular Engineering

Georgia Institute of Technology
December 2012

DEVELOPMENT OF AN AMINE DEHYDROGENASE

Approved by:

Dr. Andreas S. Bommarius, Advisor
School of Chemical & Biomolecular Engineering
Georgia Institute of Technology

Dr. Jeffrey Skolnick
School of Biology
Georgia Institute of Technology

Dr. Christopher W. Jones
School of Chemical & Biomolecular Engineering
Georgia Institute of Technology

Dr. John W. Wong
Biocatalysis Center of Emphasis
Chemical Research & Development
Pfizer Global Research & Development

Dr. Yoshiaki Kawajiri
School of Chemical & Biomolecular Engineering
Georgia Institute of Technology

Date Approved: August 13, 2012

To my parents, Joseph & Deborah

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my parents Joseph and Deborah. Your guidance and unconditional support has been invaluable to my success throughout college. The encouragement from my entire family has been so helpful throughout graduate school.

I would also like to thank my advisor, Prof. Andreas Bommarius for his direction, instruction, and patience over the last five years. Your input and optimism has kept me 'plowing ahead' and has been instrumental in my scientific development. I would like to thank my committee members; Prof. Christopher Jones, Prof. Yoshiaki Kawajiri, Prof. Jeffrey Skolnick, and Dr. John W. Wong. Thank you all for your time, encouragement, and support.

All of the members of the Bommarius lab, past and present, have made my graduate career not only productive, but enjoyable. We have shared many great and unforgettable moments. In particular, I would like to extend my sincere gratitude to Dr. Eduardo Vazquez-Figueroa for being an outstanding mentor and enabling me learn a tremendous amount early on, in addition to his contributions to the amine dehydrogenase project. I would like to thank Dr. Janna Blum for her frequent and generous help with experiments, explanations of molecular biology, and contributions to the Current Opinions publication. Dr. Russell Vegh, thank you for helpful discussions and distraction from research when a break was need. I would like to recognize Nicolas Woodall for your years of hard work as an undergraduate. Jonathan Rubin, thank you for your extensive review of my manuscripts and many fruitful discussions. Michael Rood and

Ryan Clairmont, thank you for your help in gathering and interpreting of NMR data. Dr. Melanié Hall, Dr. Yanto Yanto, Dr. Thomas Rogers, Dr. Prabuddha Bansal and Jonathan Park, your advice and friendship has been greatly appreciated. Dr. Bettina Bommarius, thank you for your help with stubborn PCRs. Samantha Au, thank you for all your hard work on the project.

I would like to thank the following funding organizations for their generous support; the Georgia Institute of Technology President's Fellowship, the Center for Pharmaceutical Development (and accompanied industrial sponsors), and the U.S. Department of Education's Graduate Assistance in Areas of National Need.

Additionally, I would like to thank the Georgia Institute of Technology, School of Chemical and Biomolecular Engineering for this incredible opportunity. The Parker H. Petit Institute for Bioengineering and Bioscience, thank you for providing a marvelous working atmosphere. I would like to thank Dr. Nicholas Hud and the members of the Hud Lab for their training and use of the CD.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

Å	angstroms
ACS	American Chemical Society
AmDH	amine dehydrogenase
API	active pharmaceutical intermediate
ATCC	American Type Culture Collection
B-PER	bacterial protein extraction reagent
bp	base pairs
CASTing	combinatorial active-site saturation testing
CD	circular dichroism
DMBA	dimethyl butyl amine
ε_{340}	extinction coefficient at 340 nm
E	average disruption
E.C.	Enzyme Commission
E. coli	*Escherichia coli*
e.e.	enantiomeric excess
EPC	enantiomerically pure compounds
epPCR	error prone PCR
FACS	fluorescence activated cell sorting
FDA	Food and Drug Administration
FDH	formate dehydrogenase
GCI	Green Chemistry Institute
GDH	glucose dehydrogenase
GluDH	glutamate dehydrogenase

IMAC	immobilized metal ion affinity chromatography
INT	2-(4-iodophenyl)-3-(4-nitrophenyl)-5-phenyltetrazolium chloride hydrate
ISM	iterative saturation mutagenesis
IPTG	Isopropyl β-D-1-thiogalactopyranoside
k_{cat}	catalytic rate constant
k_{uncat}	uncatalyzed rate constant
K_M	Michaelis constant
kDa	kilodaltons
LB	Luria Bertani
LB-AMP	Luria Bertani containing ampicillin
LeuDH	leucine dehydrogenase
LeuDH-AmDH	LeuDH-based amine dehydrogenase
MBA	methyl benzyl amine
MIBK	methyl isobutyl ketone
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
Ni-NTA	nickel-nitriloacetic acid
NMR	nuclear magnetic resonance
OD_{600}	optical density at 600 nm
PCR	polymerase chain reaction
PDB	Protein Data Bank
PES	phenazine ethosulfate
PFPA	*para*-fluoro phenyl acetone
PheDH	phenylalanine dehydrogenase
PheDH-AmDH	PheDH-based amine dehydrogenase
ProSAR	PROtein Sequence Activity Relationships

RFU	relative fluorescence units
RMSD	root-mean-square deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
$t_{1/2}$	half life
T_M	melting temperature
TA	transaminase
TFAA	trifluoroacetic anhydride
TTN	total turnover number
ValDH	valine dehydrogenase
ValDH-AmDH	ValDH-based amine dehydrogenase
WT	wild type

SUMMARY

Biocatalysts are increasingly prevalent in the large-scale synthesis of enantiomerically pure compounds, which are mainly used as active pharmaceutical ingredients. Enantiomerically pure forms can lead to lower dosages, improved efficacy, and even allow for extension of patents. However, many sought-after reactions lack a suitable enzymatic production route. A recent assessment by the ACS Green Chemistry Institute, Pharmaceutical Roundtable noted that the asymmetric synthesis of amines from prochiral ketones and free ammonia was one of the top aspirational reactions challenging the pharmaceutical industry.

The first goal of this work was developing an amine dehydrogenase through the application of directed evolution altering the substrate specificity of an existing leucine dehydrogenase scaffold. The resulting amine dehydrogenase asymmetrically would convert methyl isobutyl ketone to 1, 3-dimethyl butyl amine. Eleven rounds of directed evolution completely altered the enzyme's specificity and successfully created amination activity. These variants were screened by various high-throughput assays identifying minute increases in amination activity, and successful mutations were carried into future rounds of mutagenesis. The resulting quadruple mutant, K67S/E113V/N261L/V290C exhibited novel reductive amination activity of 0.69 U mg^{-1} with a corresponding kcat value of 0.46 s^{-1}. The enantioselectivity of the wild-type enzyme was maintained despite the drastic changes to the binding pocket and yielded (*R*)-1,3-DMBA with an *e.e.* value of 99.8% at 92.5% conversion making it an attractive catalyst in the synthesis of chiral amines. This was the first example of a cofactor-dependent amine dehydrogenase capable of selectively synthesizing chiral amines from a prochiral ketone and free ammonia.

The second goal was to develop another amine dehydrogenase by applying these previously identified, influential mutations to a phenylalanine dehydrogenase scaffold. A single library of the two most influential residues was screened by a stringent absorbance-based assay. The resulting K77S/N276L double variant enhanced the desired reductive amination activity to a k_{cat} value of 6.85 s^{-1} at 25 °C. The high enantioselectivity was maintained with an *e.e.* value of > 99.8% toward the (*R*)-enantiomer. When paired with a cofactor recycle system, conversions in excess of 90% were achieved while maintaining selectivity. This second amine dehydrogenase exhibits significantly enhanced activity and different substrate specificity compared to the previous enzyme.

The final goal of this work was to test the applicability of the mutations responsible for amine dehydrogenase activity across related scaffolds in the enzyme's sub-family. During the evolutionary route of producing the aforementioned enzymes, the evolutionary process concurrently identified two mutations as essential for amination activity. The mutation of these residues is believed to be broadly applicable across the Glu-Leu-Phe-Val dehydrogenase sub-family of enzymes. To test this hypothesis, a third scaffold from valine dehydrogenase will be evolved at its analogous residues for amine dehydrogenase activity, potentially giving credence to a novel enzyme family.

Lastly, future works are discussed exploring the applicability of this enzyme to target substrates, conversion in organic media, and scale up of the reaction.

CHAPTER 1

INTRODUCTION

Enzymes are unique in their ability to significantly enhance reaction rates under mild conditions with high levels of efficiency and selectivity. After substrates have bound to an enzyme's active site, their half-life is usually only a fraction of a second (1). Some enzymes perform catalysis so well they border on the physical limitations of efficiency with encountering substrate in solution, and have been known to achieve rate enhancements (k_{cat}/k_{uncat}) in excess of 19 orders of magnitude. These outstanding examples of enzymatic catalysis establish their remarkable power, and demonstrate their potential use in chemical synthesis.

In addition to rate enhancement, many enzymes are selective in their catalysis and allow for the direct production of single enantiomers. Single enantiomers of target product molecules have a large importance in the pharmaceutical and agricultural industries. Therapeutic compounds often act as structurally optimized inhibitors of biological processes. Since the human body functions using chiral chemistry, these compounds almost always contain chiral centers. In 2006, 80% of small-molecule drugs approved by FDA were chiral and 75% were single enantiomers (2). Current USA Food and Drug Administration (FDA) regulations require proof that any non-therapeutic isomer comprising over 1% of the total composition be non-teratogenic. Thus, a racemic drug would require separate toxicology studies of each enantiomer. Further, as the complexity of these compounds increases, molecules often require more than one chiral

center (3). It is common for pharmaceutical processes to require a minimum acceptable enantiomeric excess (*e.e.*) of 98% to circumvent these difficulties.

Chiral intermediates are classically prepared by three different routes (4). One method is to simply attain them through a naturally occurring chiral synthon. This allows for an inexpensive and readily available source of natural products, but this pool of compounds is limited. Alternatively, chirality can be achieved through the resolution of a racemic mixture, as seen in the kinetic resolution of alcohols and amines through the application of lipases or esterases (5-7). The major disadvantage of kinetic resolution is an unfavorable equilibrium constant, usually near unity. The use of a second reaction is required to drive the reaction to completion. Lastly, the chiral centers can be synthesized directly using an asymmetric catalyst or enzyme. If the desired selectivity is not achieved initially, multiple enrichments, such as crystallization, can be required to achieve an increased selectivity. In comparison to other catalysts, enzymes usually offer superior enantio- and regioselectivity, commonly reporting *e.e.*'s of > 99.9% (8). High initial selectivity eliminates the need for enrichment, significantly reducing the cost of production.

One of the short comings of biocatalysis is the limited number of reactions which have identified enzymatic routes. In 2005, the American Chemical Society's (ACS) Green Chemistry Institute (GCI) and several leading global pharmaceutical companies created the ACS GCI Pharmaceutical Roundtable. The goals of this center encompassed innovation in the discovery, development, and production of pharmaceuticals. Consequently, the Roundtable identified the most aspirational reactions currently challenging the pharmaceutical industry (Table 1.1) (9). The reductive amination of a

prochiral ketone with free ammonia to produce chiral amines ranked second on this list, and had yet to be achieved.

Table 1.1. Aspirational reactions challenging the pharmaceutical industry as determined by the ACS GCI Pharmaceutical Roundtable (9).

Research Area	**Number of Roundtable companies voting for this research area as a priority area**
C–H activation of aromatics (cross coupling reactions avoiding the preparation of haloaromatics)	6 votes
Aldehyde or ketone + NH_3 + ''X'' to give chiral amine	**4 votes**
Asymmetric hydrogenation of unfunctionalized olefins/enamines/imines	4 votes
New greener fluorination methods	4 votes
N-Centered chemistry avoiding azides, hydrazine etc	3 votes
Asymmetric hydramination	2 votes
Green sources of electrophilic nitrogen (not TsN_3, nitroso, or diimide)	2 votes
Asymmetric hydrocyanation	2 votes

This work achieves exactly this reaction through the evolution of an amino acid dehydrogenase scaffold to an amine dehydrogenase. Developments in recombinant technology have allowed for expansion to the biocatalysis repertoire, in turn enabling new chemistries. Directed evolution has emerged as a common tool in the creation of novel enzymes. Directed evolution involves the introduction of mutations to a parent enzyme, followed by screening or selection of these variants for a desired property, and propagation of these beneficial mutations into subsequent cycles of mutagenesis to achieve a desired property.

To apply this process, a robust and sensitive high-throughput assay has been developed allowing identification of favorable mutations. Eleven rounds of mutagenesis had completely altered the substrate specificity resulting in the asymmetric amination of methyl isobutyl ketone to (*R*)-1, 3-dimethyl butyl amine, while maintaining high selectivity and conversion. The influential mutations were identified and inserted into an analogous scaffold of phenylalanine dehydrogenase, further increasing the substrate specificity and activity of AmDHs. These mutations will lead to the possible creation of a family of AmDHs.

CHAPTER 2

STATUS OF PROTEIN ENGINEERING FOR BIOCATALYSTS: HOW TO DESIGN AN INDUSTRIALLY USEFUL BIOCATALYST

2.1 Introduction

Recent advances in the development of both experimental and computational protein engineering tools have enabled a number of further successes in the development of biocatalysts ready for large-scale applications. Key tools are first, the targeting of libraries, leading to far smaller but more useful libraries than in the past, second, the combination of structural, mechanistic, and sequence-based knowledge often based on prior successful cases, and third, the advent of sequence and structural based algorithms allowing the design of novel functions (10-13). Based on these tools, a number of improved biocatalysts for pharmaceutical applications have been presented, such as an (*R*)-transaminase for the synthesis of active pharmaceutical ingredients (APIs) of sitagliptin (Januvia®) and ketoreductases, glucose dehydrogenases, and haloalkane dehalogenases for the API synthesis toward atorvastatin (Lipitor®) and montelukast (Singulair®) (14-18).

The ideal industrial biocatalyst should have high specific activity combined with high specificity on the reactant in question, the highest possible enantioselectivity in case of generating a chiral center, and should be stable at industrial process conditions, that is, often higher than ambient temperature and in partially or wholly nonaqueous solvents. The advent of molecular biology since the late 1970's embodied in protein engineering protocols; progress in crystallography and computational methods has enabled the design

of protein catalysts for targeted applications and improvement of desired traits. When starting any protein engineering project, two questions must be asked, 'what template should I start with?' and 'what tools should I use from the biocatalyst toolbox?' Protein engineering has evolved from its first phase of rational design to the second phase with combinatorial design or 'directed evolution', and most recently to the third phase, data-driven design of biocatalysts. No longer is a large library size the driving attribute for protein development but instead a small number of targeted libraries, which embody a restricted range of codons and predictive tools optimizing the trait in question.

This chapter covers the progress in protein engineering over the last few years and describes status and perspectives for biocatalyst design and optimization.

2.2 Protein engineering tools

Numerous protein engineering tools have emerged in recent years for the improvement of existing biocatalysts or for their adaptation to novel substrates. Table 2.1 lists a number of different enzymes that have been evolved over the past few years along with the key results from these protein engineering efforts. The selection of protein engineering tools depends upon the following criteria:

(i) The expected involvement of amino acid residues close or distant to the active site: changes in activity or selectivity often involve residues close to the active site, where improvements in stability in most cases require targeting residues away from the active site (19).

The library screening capacity: whereas combinatorial design strove for the largest possible library sizes, topped by selection (growth-based selection: up to 10^{12} variants) and routinely hitting 10^{5}–10^{7} variants for

screening with fluorescent (such as fluorescence-activated cell sorting, FACS (20)) or spectrophotometric assays, data-driven protein engineering strives for the smallest possible library, usually lower than 100,with the techniques described below.

(iii) The depth and accuracy of available sequence and structure information: crystal structures accurate to < 1 Å allow positioning of water, often critical to the catalytic mechanism (21, 22), while many data-driven concepts (see below) depend on the availability of as many as possible sequences with the maximum possible level of identity for accurate predictions.

A description of individual tools follows.

Table 2.1. Selected results from protein engineering of biocatalyst (23).

Enzyme	Protein engineering goal	Tools	Metric(s)	Ref.
ATA-117 D-amino acid transferase	Overall enzyme redesign, substrate specificity, thermostability, organostability	Targeted restricted codon libraries around active-site, ProSAR, directed evolution	200 g/L substrate converted with >99.95% ee, 92% assay yield obtained in ~24 h	(17)
Halohydrin dehalogenase	Activity	ProSAR	Volumetric productivity - 100 g L^{-1} d^{-1} (4000-fold improvement)	(10)
Ketoreductase	Organostability, enantioselectivity	ProSAR	>100 g L^{-1} substrate converted with >99% ee (template had 63% ee), >99% assay yield obtained in ~24 h	(24)
Glucose dehydrogenase	Improved thermostability and organostability	Structure-guided consensus	$t_{1/2}$ > 3.5 days at 65 °C (x 10^6 fold improvement)	(25, 26)
Ketoreductase-glucose dehydrogenase	Activity and stability	Directed evolution, DNA shuffling	160 g L^{-1} substrate converted with >99.9% ee, 95% isolated yield obtained in ~8 h	(27)
Kemp eliminase	Novel catalysis	*De novo* design via Rosetta, directed evolution	k_{cat} = 1.375 s^{-1}, k_{cat}/K_M = 2590 M^{-1} s^{-1}	(28)
Nucleoside kinase	Altered substrate specificity	Rosetta molecular modeling	Relative specificity $((k_{cat}/K_M)_{ddT}/(k_{cat}/K_M)_T)$= 8.5 (8500-fold improvement)	(29)
Cytochrome P450 (P450-BM)	Regioselectivity	Site-saturation mutagenesis based on docking knowledge	Indirubin/indigo ratio = 9 (51-fold improvement)	(30)
Epoxide hydrolase	Enantioselectivity	Iterative CASTing	E = 115 (25-fold improvement)	(31)
Cellobiohydrolase class II	Improved thermostability	SCHEMA	$t_{1/2}$ ~3200 min at 63 °C (34-fold improvement)	(32, 33)
Diels-Alderase	Novel catalysis	*De novo* design via Rosetta	k_{cat} = 0.036 s^{-1}; k_{cat}/K_M = 0.455 M^{-1} s^{-1}	(12)
Cellobiohydrolase class I	Improved thermostability	SCHEMA	T_{50} = 65.7 °C, increased 3.4 °C compared to parent	(34)

2.2.1 Iterative saturation mutagenesis

Iterative saturation mutagenesis (ISM) has become known as an essential tool in designing novel biocatalysts (35). The saturation of individual residues has been proven to create novel functionality, enantioselectivity, and thermostability. The incremental accumulation of single site mutations likely mirrors the evolutionary path of nature (36), although this stepwise procedure is not the most direct method and will likely miss any synergistic effects from neighboring residues (37).

2.2.2 Combinatorial active-site saturation testing

Combinatorial active-site saturation testing (CASTing) is another powerful tool for developing more active, thermostable, or selective mutants (35). CASTing's usefulness lies in its ability to test the synergistic effects of saturating small portions of the protein's active site. However, CASTing libraries can easily exceed screening capacity because of the exponential growth of the screening requirement with each additional mutation. A four position NNK library would result in 32^4 codons and a screening requirement of $> 10^6$ colonies. CASTing is best applied when activity toward the desired substrate already exists at some level, otherwise low levels of activity may be missed depending on the precision of the screening assay (Chapters 3-5 are an exception to the strategy). Despite their low activity, these transitional mutants are essential in the iterative process to a successful catalyst. The B-FIT approach (35, 38), a variation of CASTing based on the temperature dependence of spatial flexibility of residues (B-factor), targets amino acid positions based on high crystallographic B-factors indicating high flexibility associated with positions in the crystal structure to improve

thermostability. This method successfully improved a lipase (lipA) but needs further corroboration with other examples (35).

2.2.3 Restricted libraries

One of the simplest ways to improve screening efficiency is the appropriate application of degenerate codons. Degenerate codons can be used to limit mutation of target sites to a desirable group of amino acids. Selection of the allowed mutations can be restricted based on rational design (39, 40) or broad mutation to efficient codons, such as NDT. While NDT limits the mutation to 12 diverse amino acids (from 12 codons), it removes the redundancy of analogous codons seen in conventional NNK codon degeneracy (20 amino acids/32 codons, including a stop codon) (41, 42). Further restriction of degenerate codons can be used to reduce library size even further (43) and was successfully used by Jochens et al. by restricting codon usage based on 3DM analysis of the α/β hydrolase fold family (41, 44).

2.2.4 Structure-guided consensus

The structure-guided consensus technique combines sequence-based and structural data and employs a set of phylogenetically diverse but functionally proven proteins. The resulting alignment of sequences often has a relatively low level of sequence identity (40–80%) (13). From this alignment, mutational points-of-interest are limited to those positions which match the desired consensus cut-off. Further constriction of the remaining residues is achieved via structural criteria, such as keeping a distance from the active site (often 4-8 Å), introducing helix stabilizers, and not breaking salt-bridge formation. The remaining consensus mutants are expressed and screened as single and multiple residue mutants for thermostability. The result is a highly constrained

mutational space, which has been shown to contain an increased percentage of thermostable mutants.

Lately, significant advances in protein engineering have come through computational methods. Computational design greatly decreases the need of probing randomized sequence space, rendering the route to novel biocatalyst much more efficient. Several significant approaches to rational design have been introduced over the last few years. The most prominent of these methods are SCHEMA, ProSAR, and ROSETTA.

2.2.5 SCHEMA

SCHEMA is a computational algorithm which estimates the amount of disruption to a protein structure upon caused DNA recombination. Libraries of chimeras are scored (E = average disruption) according to the number of residue–residue contacts which broken when compared to the parent proteins. Screening can then be focused to those chimeras with less structural disruption (low E). The SCHEMA algorithm ultimately results in an enhanced probability of functional chimeras with relatively low identity to the parent sequences, effectively generating a family of diverse, folded, and functional proteins (45). Additionally, many nonfunctional chimeras with low E have regained activity by low error-rate random mutagenesis (11).

2.2.6 ProSAR

Richard Fox et al. have developed a strategy for improving recombination-based directed evolution by applying statistical analysis of PROtein Sequence Activity Relationships (ProSARs) (10). After initially screening a diverse group of mutants for activity toward the desired function, this data is applied to the ProSAR model as a training set. The ProSAR model sorts mutations based on their predicted effect on

activity. Once beneficial mutations are identified, they are fixed for subsequent rounds. After each iteration, the resulting mutants showed constant improvement over the parent protein.

2.2.7 Rosetta

Currently the Rosetta methodology (12, 28, 46) is the most rational of the computational techniques. Rosetta applies the quantum mechanical and computational design of the active site around a reaction's transition state to create a novel biocatalyst from an existing template structure. This method has been shown to create Diels-Alderase and previously nonexistent Kemp eliminase enzymes (12, 28, 47, 48). Despite exhibiting relatively low catalytic activity, they achieve some level of base activity to be improved upon (46). As protein structure–function knowledge advances to incorporate further interactions, purely rational techniques such as Rosetta will create increasingly robust enzymes.

2.3 Protein engineering challenges

One of the most difficult problems in protein engineering is obtaining new function with an enzyme that has no detectable initial activity of the desired substrate. Peisajovich and Tawfik suggested that the commonly cited first rule of directed evolution 'you get what you select/screen for' should be followed by a second rule of 'you should select for what is already there' (49). While there are exceptions to every rule, obtaining that initial activity does seem to be the key to further success. Gerlt and Babbitt describe enzyme (re)design from a variety of different templates; however, these (re)designs still need to improve to reach activity levels required for industrial catalysts, with k_{cat}'s ranging between 0.000007 and 3.9 s^{-1} (50). Random mutagenesis has been the most

common approach to tackling this problem, it is of interest to move toward more rational approaches to decrease screening effort and increase understanding (51). An algorithm on how to take a poor, but functional catalyst, to industry has yet to be laid out (Figure 2.1).

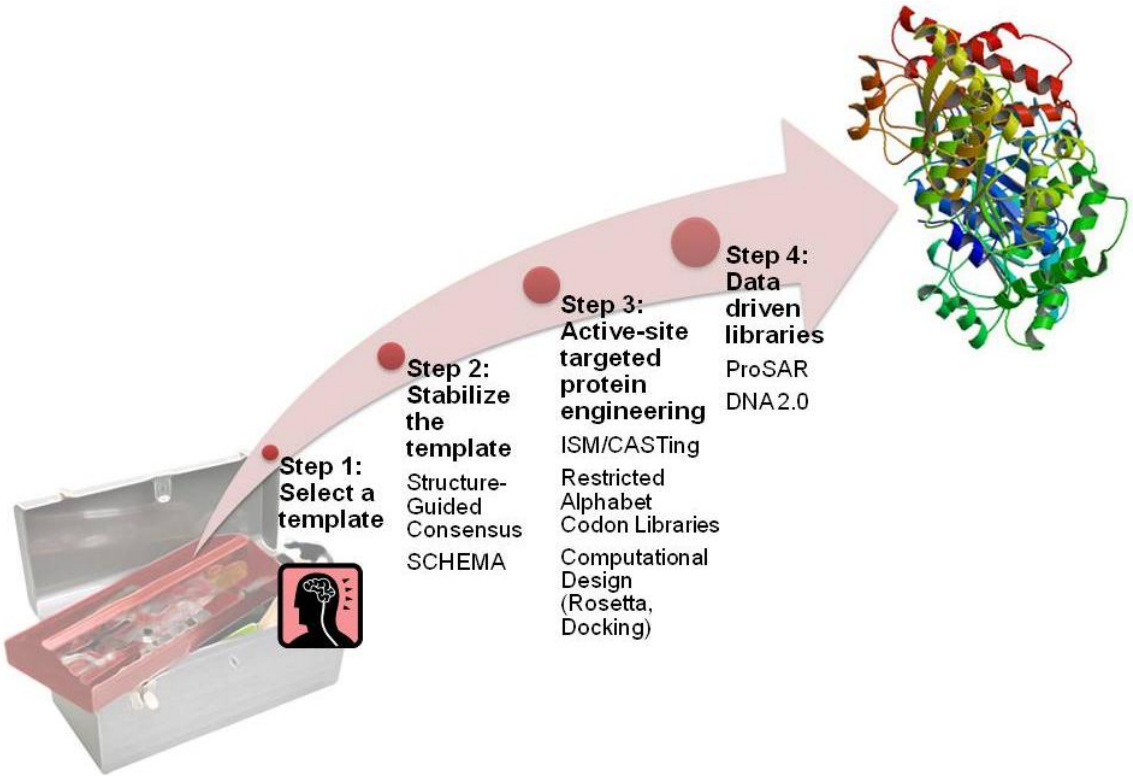


Figure 2.1. Protein engineering toolbox and design algorithm toward an industrial catalyst (23).

2.3.1 Proposed algorithm to design an industrial biocatalyst

Step 1: select the template

The question of what template to start with often puzzles the researcher and is not often discussed in the literature, as most papers just start with a statement of which protein was chosen. Should one start with a more ancestral protein with many

promiscuous (but low) activities?, a very stable template?, a highly active protein that has poor selectivity?, and an easily expressible protein with high expression levels? Often these criteria are mutually exclusive, so how does one prioritize?

We propose that a template that has the most desirable reaction properties should be selected as the starting point, preferably one with some activity toward the target substrate. Clearly, this is not possible when no class of enzymes exists for the desired reaction, in which case one would have to apply the de novo approaches. One of the best methods of selecting a template enzyme has been developed by Höhne et al. and is based on an *in silico* analysis of multiple sequences, structures, and motifs that led straight to selection of a template with desirable properties without necessarily requiring Steps 2–4 (52, 53).

Step 2: stabilize the template

Thermostability, often closely related to organic solvent stability (25, 54), is an important parameter in the development of industrially applicable biocatalysts. Improved thermostability increases the life of the enzyme and the number of total turnovers that the enzyme can perform over the life of the catalyst, known as total turnover number (TTN). Several different protein engineering techniques have been successfully used to improve protein thermostability, including error-prone PCR (epPCR) (55), DNA shuffling via SCHEMA (11, 32, 33, 45, 56), temperature value (B-FIT) approach (35, 38) and the structure-guided consensus concept (26), all of which except epPCR require additional structural information. In addition to improved enzyme life, thermostability has been linked to the ability of an enzyme to tolerate destabilizing mutations (36, 49, 57, 58). The selection of the starting template can also influence the result of gaining new catalytic

function. ‘Neutral’ drifted and thermostabilized templates have higher tolerance for mutations, as such they can tolerate considerable mutation toward new specificities. ‘Neutral’ drift is achieved by introducing random mutations (such as epPCR or DNA shuffling) and selecting those variants for wild-type activity (49, 57, 58). Thus, ‘neutral’ mutations are accumulated in the library in and around the active-site that may increase promiscuity. The ‘neutral’ drift method is highly dependent on an effective high-throughput screen or selection assay to obtain the library of neutral mutations. However, it is very straightforward to create a consensus sequence and one does not require any high-throughput selection to do so. The TEM-1 β-lactamase neutral drift evolution by Bershtein et al. demonstrated that all stabilizing mutations discovered using directed evolution drifted back to the consensus sequence (57). Therefore, moving toward a thermostabilized template, via the consensus sequence as described in section ‘Protein engineering tools’, before introducing libraries toward altered substrate specificity will increase the number of active variants and have the highest degree of success.

Step 3: semi-rational design around the active-site

To design toward altered substrate specificity or enantioselectivity, mutations around the active site should be chosen (19). Rational design using the tools described in section ‘Protein engineering tools’, in particular CASTing, ISM and restricted codon libraries.

Step 4: machine-based learning via recombination

Lastly, machine-based learning, such as ProSAR, remains the most effective way to improve an enzyme with small library sizes.

2.4 Successful case studies

One of the most successful industrial biocatalyst redesigns was the development of the ATA-117 D-amino acid amino transferase for the Januvia™ (Sitagliptin phosphate) process. A truncated substrate was used to overcome the lack of activity on their target substrate (17). The wild-type enzyme had detectable activity (albeit low ~4% conversion after 24 hr) toward the truncated substrate. Savile et al. were able to use semi-rational design via transition state docking to improve the activity toward the truncated substrate, resulting in an enzyme capable of aminating the target substrate at low conversion (< 1% conversion after 24 hr). After several more rounds of targeted and random mutagenesis via ProSAR, an enzyme designed for industrial yields and operating conditions was developed. Ultimately, the final enzyme contains 27 mutations both around the active-site and at the dimer interface.

Approaches that also target the active site such as CAST-ing along with structural information from docking studies have generated some very well developed biocatalysts. Iterative CASTing and ISM have also been successfully used toward improved enantioselectivity (31). This technique was most recently used to generate enoate reductases mutants that were both (*R*)- and (*S*)-selective (37). Hu et al. discovered a cytochrome P450 BM-3 mutant that altered the regioselectivity of the P450 BM-3 toward indirubin through the evaluation of a single NNK library that was designed based on docking studies (30).

Codexis has successfully applied ProSAR (in conjunction with directed evolution) to enhance the activity and enantioselectivity of several keto-reductases toward industrially relevant substrates. During this process, the resulting catalysts are also

optimized toward the desired industrial reaction conditions (14, 24, 27, 59). Other successful approaches include the computational design of active sites. The Baker lab has been successful using the Rosetta algorithm to design de novo biocatalysts including a Kemp elimination enzyme (46, 60) and a Diels-Alderase (12). These catalysts challenge our understanding on structure–function relationships and significantly broaden our ability to design catalysts not found in nature. Though, this technique requires significant improvement, as the catalytic rates are slow relative to naturally occurring enzymes, owing to the lack of malleability (i.e. excessive stiffness) of the protein binding site and active-site epitope and thus far the catalytic constants k_{cat} are mostly < 1 s^{-1}, far below what is required of a viable industrial catalyst. A significantly improved situation concerns Rosetta's application to design orthogonal nucleoside kinases: the algorithm resulted in successfully identified residues for specificity change and suggested mutations to create the concerted mechanism required for selectivity (29).

2.5 Future outlook of protein engineering of biocatalysts

The trend of improving proteins with targeted libraries incorporating structural and mechanistic knowledge as well as algorithms for interpreting data points on the function-sequence map likely will continue and even accelerate. Less dependence on large-scale screening equipment and ultimately a higher hit rate per unit of resource spent drive this development.

2.6 Conclusions

Several of the described tools, such as targeted libraries, restricted codons, or structure-guided consensus, have contributed to improve biocatalysts to levels useful in large-scale applications. Computational methods for improved reactivity based on

structures, i.e. that is, a structural picture of the transition state, seem to require further optimization to yield fast catalysts; however, they already have yielded new catalysts.

CHAPTER 3

DEVELOPMENT OF A NOVEL AMINE DEHYDROGENASE FROM LEUCINE DEHYDROGENASE

3.1 Introduction

Biocatalysts are increasingly prevalent in the large-scale synthesis of enantiomerically pure compounds (EPCs), which are mainly used as active pharmaceutical ingredients (APIs). Enantiomerically pure forms can lead to lower dosages, improved efficacy, and even allow for extension of patents (61). In 2006, 80% of small-molecule drugs approved by FDA were chiral and 75% were single enantiomers (2). However, many sought-after reactions lack a suitable enzymatic production route. A recent assessment by the ACS Green Chemistry Institute, Pharmaceutical Roundtable noted that the asymmetric synthesis of amines from prochiral ketones and free ammonia was one of the top aspirational reactions challenging the pharmaceutical industry (Figure 1.1) (9).

Our novel enzyme achieves exactly that aim: it creates amines with high selectivity. Previously characterized amine dehydrogenases were incapable of effecting the reductive amination of ketones and lacked stereoselectivity (62).

While chiral amines can be produced both chemically and enzymatically, the large-scale production of chiral amines is still challenging and heavily reliant on traditional methods of chemical synthesis (27, 61, 63-65). Common methods include resolution though fractional crystallization (66) and the hydrogenation of C=N bonds,

particularly in enamines (67). Nonetheless, some chemoenzymatic routes, particularly with transaminases (47, 68-71), have shown promise in the dynamic kinetic resolution of racemic amines (70, 72, 73) and the direct asymmetric synthesis of amines with ω-transaminases (ω-TA), as used in the synthesis of sitagliptin (74-76). This novel process has eliminated the use and removal of a less-selective rhodium catalyst, yet requires the use of a sacrificial amine source. The undesired ketone by by-product must also be removed to shift the reaction equilibrium beyond about 50% conversion.

Asymmetric synthesis by amine dehydrogenases (AmDHs) would be the ideal route to produce chiral amines. When paired with a cofactor recycling system, such as glucose/glucose dehydrogenase (GDH) or formate/formate dehydrogenase (FDH), amine dehydrogenases allow for the direct production of chiral amines, with the consumption of only an inexpensive reducing agent, such as glucose or formate, and free ammonia (Figure 3.1) (13, 27, 77, 78).

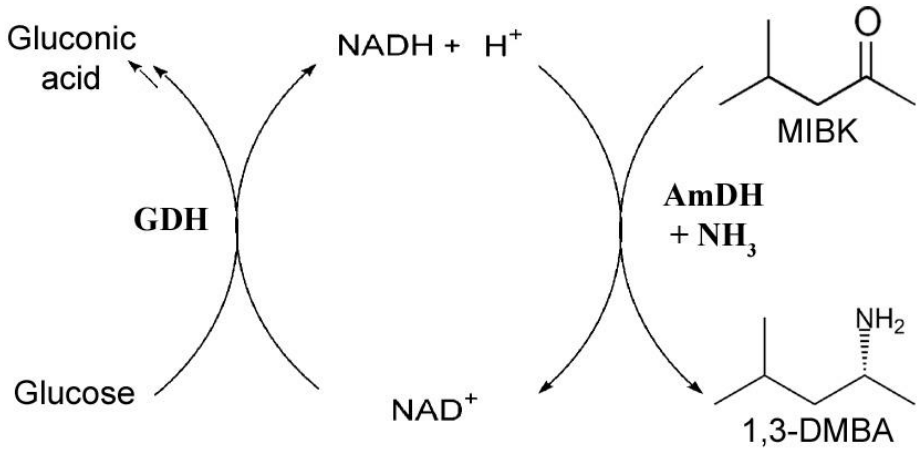


Figure 3.1. Asymmetric synthesis reaction scheme with an amine dehydrogenase paired with a glucose dehydrogenase cofactor recycling system (79).

By using an existing amino acid dehydrogenase scaffold, we have successfully altered the substrate specificity through several rounds of protein engineering to create an amine dehydrogenase. Instead of the wild-type α-keto acid, the amine dehydrogenase now accepts the analogous ketone, methyl isobutyl ketone (MIBK), which corresponds to replalcement of the carboxyl moiety with a methyl group (Figure 3.2). The wild-type leucine dehydrogenase exhibited no measurable activity toward the reductive amination of MIBK.

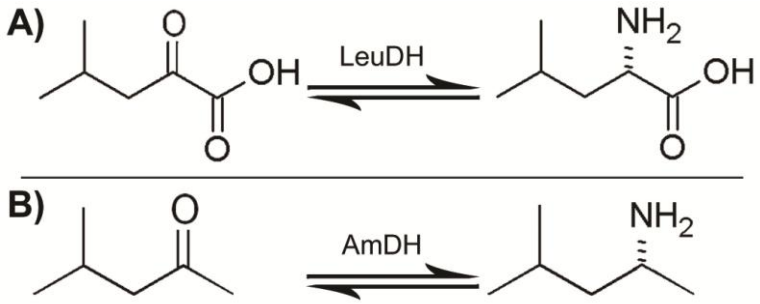


Figure 3.2. Reactions of wild type LeuDH and LeuDH-AmDH. A) Wild-type leucine dehydrogenase reaction. B) Novel amine dehydrogenase reaction (79).

3.2 Materials and Methods

3.2.1 gDNA preparation, gene isolation and overexpression

Leucine dehydrogenase (E.C. 1.4.1.9) from *Bacillus stearothermophilus* was donated by Assistant Professor Bert C. Lampson from East Tennessee State University, followed by isolation of the genomic DNA through application of method B described by Mehling et al. (80). The isolated gene (1104 bp) was inserted into pET17b vector (Life Technologies-Invitrogen, Grand Island, NY, USA) at the restriction sites *Nde*I and *Hind*III (New England Biolabs) using T4 DNA Ligase (Roche, South San Francisco, CA,

USA) at 16 °C, 8 hours. This allowed for overexpression upon transformation into *E. coli* BL21 DE3 competent cells (Life Technologies). Cultures were expressed at 37° C in LB media for 24 hours. Cultures were induced with IPTG at and OD_{600} between 0.4 and 0.6 absorbance. Overexpression was evident by SDS-PAGE identifying a high protein concentration at 40.5 kDa (Figure 3.3).

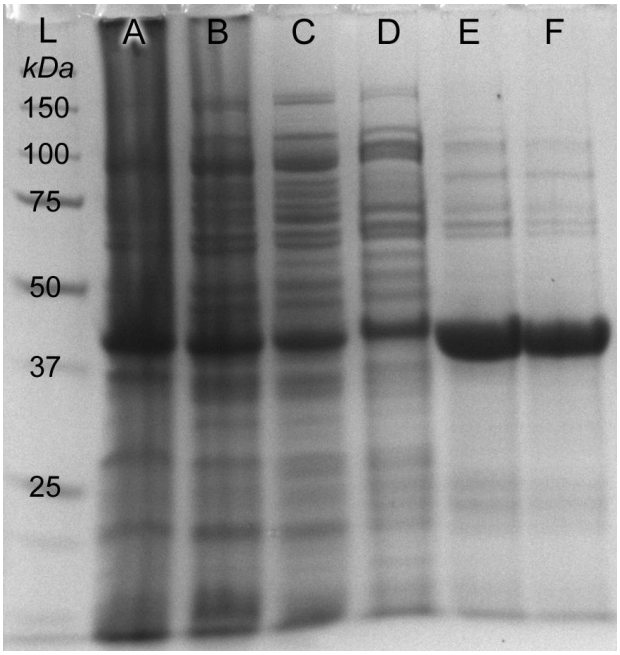


Figure 3.3. His-tag purification of wild-type LeuDH (42kDa). Lane L: Ladder, A: Clarified cell lysate, B: Unbound clarified cell lysate, C: 50 mM imidazole wash 1, D: 50 mM imidazole wash 2, E: Pure protein elution fraction 1, F: Pure protein elution fraction 2.

3.2.2 Protein purification

For purification, proteins were expressed in pET28a vector (Life Technologies) in BL21 (DE3) with an N-terminal Histag. The gene was inserted in the vector at restriction sites *Nde*I and *Hind*III. His-tagged proteins were expressed in a similar manor as (Section

3.2.1) at 37° C for 24 hours. His-tagged proteins were subsequently purified using immobilized metal ion affinity chromatography (IMAC) with a Ni-NTA resin (Thermo Scientific, Rockford, IL, USA). Cell culture pellets were made by centrifugation (4000 rpm, 30 min, 4° C) in 50 mL aliquots. Cell pellets were resuspended in 50 mM phosphate buffer (6 mL) at pH 8.0 containing 20 mM imidazole and 300 mM NaCl. This was followed by sonication and centrifugation to remove cell debris. The clarified cell lysate was bound to the Ni-NTA resin by shaking on ice for 1 hour, and subsequently purified by column affinity chromatography. Protein-bound resin and clarified cell extract was poured over a column, and washed twice with 50 mm phosphate buffer (5 mL) at pH 8.0 containing 50 mM imidazole and 300 mM NaCl. Purified protein was eluted with the same buffer solution containing 250 mM imidazole (2 mL). An SDS-PAGE gel of a representative purification can be seen below. The leucine dehydrogenase protein can be seen at 42.5 kDa with his-tag attached (Figure 3.4).

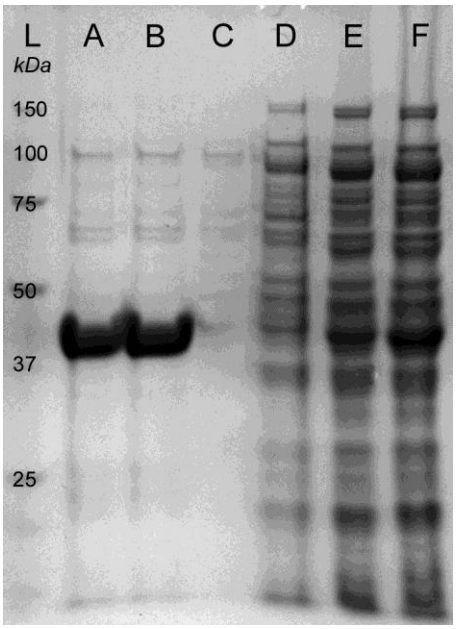


Figure 3.4. His-tag purification of K67S/E113V/N261L/V290C AmDH (42kDa). Lane L: Ladder, A: Pure protein elution fraction 2, B: Pure protein elution fraction 1, C: 50 mM imidazole wash 2, D: 50 mM imidazole wash 1, E: Unbound clarified cell lysate, F: Clarified cell lysate.

3.2.3 Spectrophotometric assay

Activity of purified proteins was measured using a spectrophotometric assay at 340 nm, corresponding to the cofactor nicotinamide adenine dinucleotide (NADH, ε_{340}= 6220 M^{-1} cm^{-1}) (81). For reductive amination, reactions took place in 500 mM NH_4Cl/NH_4OH at pH 9.6, with 200 μM NADH and 20 mM of the ketone substrate, unless otherwise specified. For oxidative deamination, reactions took place in 0.1 M $NaHCO_3/Na_2CO_3$ buffer at pH 10.0, with 1 mM NAD+ with 10 mM of the amine substrate of interest. All reactions were performed at 25° C unless otherwise specified. Enzyme was added in 10 μL aliquots in dilute concentrations. Enzyme concentrations varied upon activity levels with the substrate present, but were dilute enough to ensure linear conversion kinetics over the analytical time period (10 min to 30 min). Specific

activities were then calculated from the stoichiometric conversion of cofactor as determine by the change in absorbance at 340 nm over time.

3.2.4 Mutagenesis and library generation

Mutant libraries were generated using overlap extension PCR. After identification of mutational sites, primers were designed according to the guidelines of the QuikChange® Site Directed Mutagenesis Protocol (Agilent Technologies, Santa Clara, CA, USA) (82). These primers were then used in the overlap extension protocol described in Molecular Cloning: A Laboratory Manual (83). The overlap extension method has several advantages. It requires the successive amplification of DNA segments, making it easier to troubleshoot than the blind success or failure of the Quikchange® protocol. Additionally, the overlap extension method does not potentially carry over whole plasmid from the parent DNA into the final transformation. This significantly decreases the amount of non-mutated colonies present in the screened library. After mutation, the resulting mutated gene was digested using *Nde*I and *Hind*III, ligated into pET17b, and expressed as described in Section 3.2.1. A table of the primers used in amplification of each library is listed below (Table 3.1).

Table 3.1. Table of forward (Fwd) and reverse (Rev) primers used in the generation of LeuDH mutant libraries.

Library			Primer	
3	Fwd	5'	AGGGCTGAACGGCCGCTACATTACG**GBGNNTRABRTK**GGCACGACCGTTGCCGATATGGA	3'
	Rev	5'	ATCCATATCGGCAACGGTCGTGC**CMAYVTYANNCV**CCGTAATGTAGCGGCCGTTCAGCCCT	3'
4	Fwd	5'	ATTCATGACACGACGCTCGGCCCGGCG**DBWGBCDBW**ACGCGCATGTGGATGTACAATTC	3'
	Rev	5'	GAATTGTACATCCACATGCGCGT**WVHGVCWVH**CGCCGGGCCGAGCGTCGTGTCATGAAT	3'
5	Fwd	5'	GATTATGTGATCAAC**NNK**GGCGGCGTCATCAACG	3'
	Rev	5'	CGTTGATGACGCCGCC**MNN**GTTGATCACATAATC	3'
6	Fwd	5'	CAACGCCGGCGGC**DBS**ATCAAC**DBS**GCCGATGAGCTGTACGGC	3'
	Rev	5'	GCCGTACAGCTCATCGGC**SVH**GTTGAT**SVH**GCCGCCGGCGTTG	3'
7	Fwd	5'	CTGAACGGCCGCTACATTACG**NNK**GTTGACGTTGGCACGACCG	3'
	Rev	5'	CGGTCGTGCCAACGTCAAC**MNN**CGTAATGTAGCGGCCGTTCAG	3'
8	Fwd	5'	CGCCGGCGGCGTC**NNK**AACGTCGCCGATGAGCTGTA	3'
	Rev	5'	TACAGCTCATCGGCGACGTT**MNN**GACGCCGCCGGCG	3'
9	Fwd	5'	CGTCATCAACGTCGCCGAT**NNK**CTGTACGGCTACAACCGTGAACG	3'
	Rev	5'	CGTTCACGGTTGTAGCCGTACAG**MNN**ATCGGCGACGTTGATGACG	3'
10	Fwd	5'	CGCCGGCTCGGCG**NNK**AATCAGCTGAAAGAGCCGCG	3'
	Rev	5'	CGCGGCTCTTTCAGCTGATT**MNN**CGCCGAGCCGGCG	3'
11	Fwd	5'	CCCGCGGCATGACGTAC**DDK**AACGCGGCCGCCG	3'
	Rev	5'	CGGCGGCCGCGTT**MHH**GTACGTCATGCCGCGGG	3'
	Fwd	5'	CGCCGGCTCGGCG**DDK**AATCAGCTGAAAGAGCCGCG	3'
	Rev	5'	CGCGGCTCTTTCAGCTGATT**MHH**CGCCGAGCCGGCG	3'

The randomized libraries were plated on Genetix 10” x 10” in. agar plates (Molecular Devices, Sunnyvale, CA, USA) containing 200 mL of LB-agar with 50 μg/mL ampicillin, and allow to grow for 18-24 hours at 37 °C. The resulting colonies were autonomously picked using a Genetix QPix2 Colony Picker (Molecular Devices), and inoculated in 96-well microtiter plates containing 200 μL of LB media with 50 μg/mL ampicillin (LB-AMP). The plates were covered in Parafilm M (Bremis, Neenah, WI, USA) to prevent cross contamination between wells, and incubated at 37 °C, 200 rpm for 24 hours. The plates were then replicated into expression plates containing 250

μL of auto-inducing MagicMedia™ (Life Technologies) containing 50 μg/mL ampicillin. Glycerol (100 μL) was added to the LB-AMP 'master' plate, and stored at -80 °C for later sequence resurrection and identification. The expression plates were covered with Parafilm M and incubated at 37 °C for 24 hours, resulting in the overexpression of protein variants (Figure 3.5). The expression plates were centrifuged (1,000 rpm, 30 min, 4 °C), decanted, and stored at -80 °C until screened.

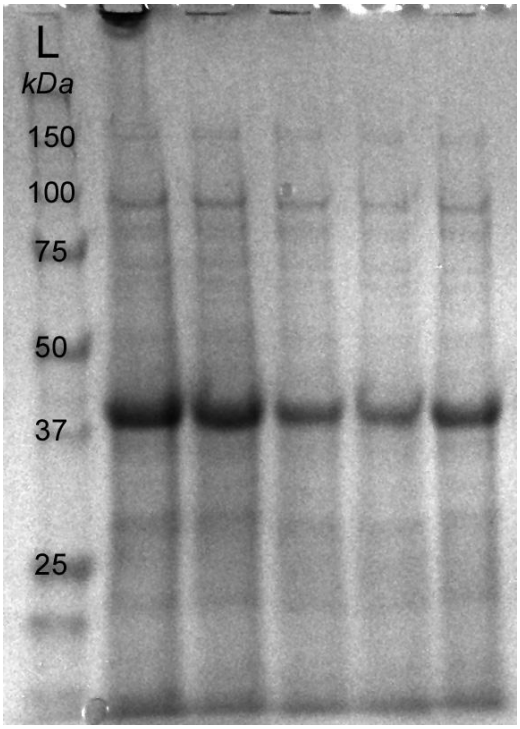


Figure 3.5. Example of library overexpression, randomly selected wells from LeuDH library 3.

3.2.5 Development of a high-throughput screening assay

Three different assays were used over the course of evaluating the mutant libraries. Each subsequent assay method improved the quality of screening over previous methods.

Libraries 2-5 were screened using an NAD+ auto-fluorescence assay (84-86). In an optimal system, sensitivities as low as 0.01 μM NAD(P)+ fluorophore were observed (87). Expression plates containing cell culture pellets were resuspended in 50 μL of B-Per Cell Lysing agent (Thermo Scientific). The resulting cell lysate was split into 20 μL aliquots in two separate plates, one reaction plate and one background plate. For the background plate, 180 μL of 1.1 mM NADH in 500 mM NH_4Cl/NH_4OH buffer pH 9.6 was added to each well. This plate lacks the ketone substrate and will only result in background conversion rates. Similarly, 180 μL of the same buffer was added to the reaction plates, containing an additional 80 mM MIBK substrate. The plates were incubated at room temperature for 2 hours to allow for conversion.

After reaction, the fluorophore was generated through a pH shift. First the residual NADH was removed by adjusting the pH to 1.0 through the addition of 30 μL of 6N HCl and incubated for 25 minutes. At this pH value, the NADH will rapidly degrade while leaving the stable NAD+. To create the NAD+ fluorescence, the pH value was re-adjust to > 13.0 by adding 80 μL of 10 N NaOH and incubated in the dark for 2.5 hours. Fluorescence was measured using a Gemini Spectramax Microplate Spectrofluorometer (Molecular Devices, Downingtown, PA, USA) at an excitation wavelength of 360 nm and emission wavelength of 455 nm (Figure 3.6) (86).

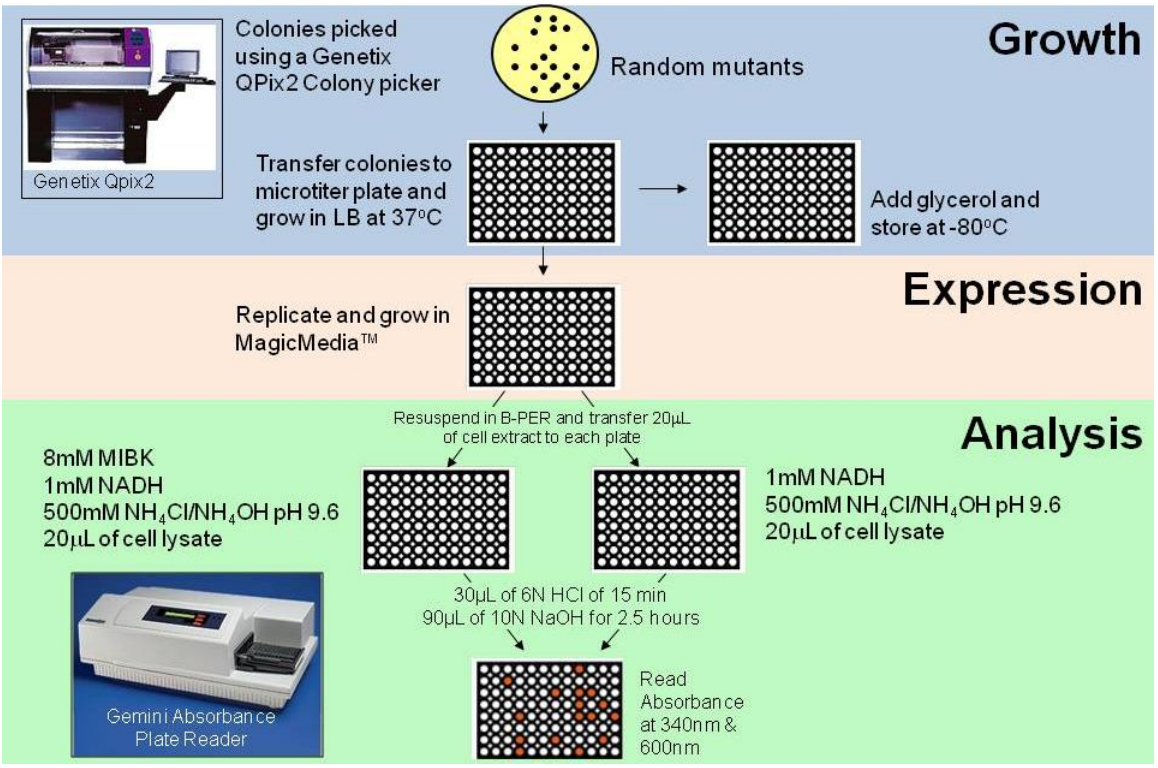


Figure 3.6. NAD+ auto-fluorescence assay used in screening libraries 2 through 5.

Wells were ranked based upon the differential increase in fluorescence over the background plate. The results of a control experiment can be seen in Figure 3.7, demonstrating simulated conversion in cell lysate by the addition of various amounts of NAD+. NAD+ concentrations can be distinguished as low as 1-2 µmol to above 100 µmol. This allows for the identification of very slight improvements in AmDH activity. The control experiment however, did not account for variations in expression levels between wells. This variation significantly reduced the sensitivity of the assay, and increased the identification of false positives. While the assay did yield successful mutants, its long reaction time and low signal to noise ratio were not ideal.

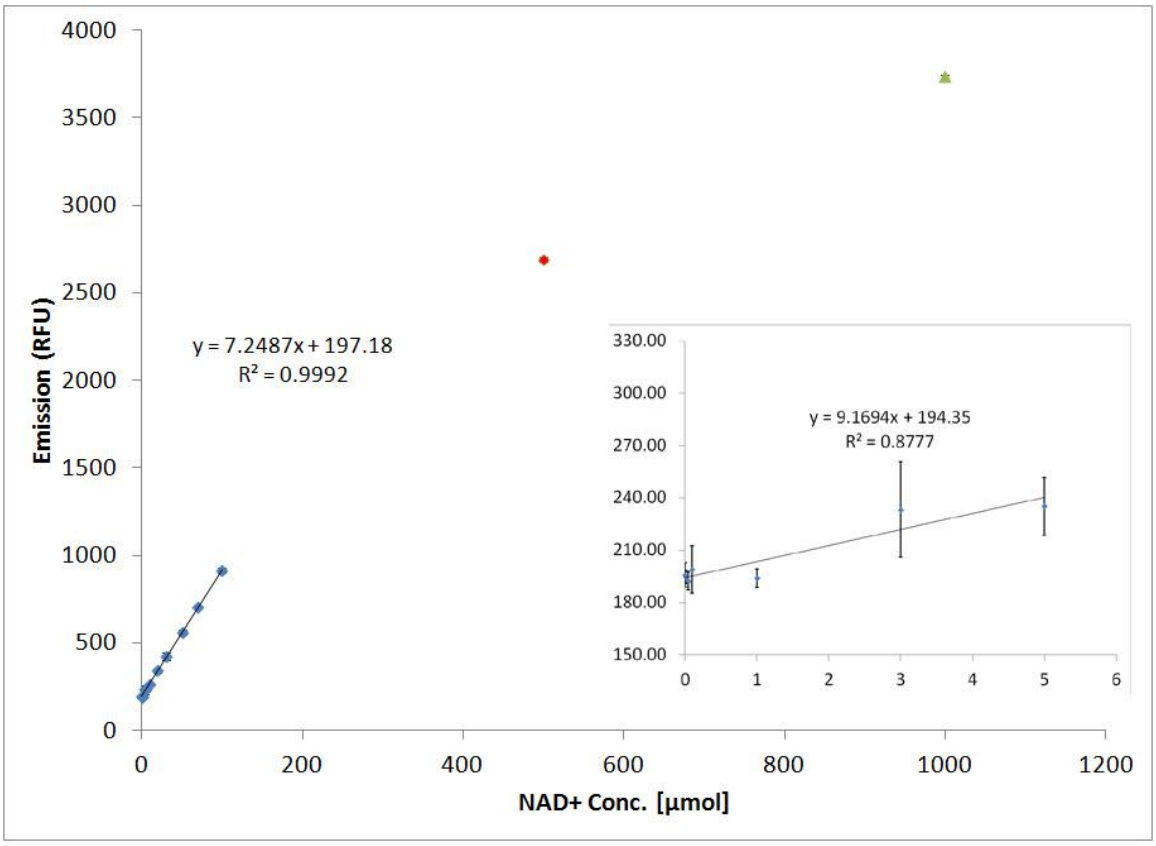


Figure 3.7. Sensitivity analysis of NAD+ auto-fluorescence assay with simulated conversion of NADH to NAD+ in BL21 cell lysate.

Libraries 6 and 7 were analyzed using a formazan-based assay developed by Chen et. al.(88). This assay takes advantage of the elevated activity in the deamination direction, giving the assay a better signal-to-noise ratio. The reaction scheme is represented in Figure 3.8. The enzymatic deamination activity is related to conversion of NAD+ cofactor to produce NADH, which is re-oxidized back to NAD+ by phenazine ethosulfate (PES, Sigma Aldrich). The resulting reduced PES subsequently reduces 2-(4-iodophenyl)-3-(4-nitrophenyl)-5-phenyltetrazolium chloride hydrate (INT, Sigma Aldrich) to create formazan. Formazan creates a deep red color which can be characterized by absorbance at 495 nm. Cell lysate was split in the same way as the

previous assay. Deamination buffer (180 µL) containing 1 mM NAD+, a catalytic amount of PES, and 0.1 M glycine buffer at pH 10.0 was added to the background plate. The reaction plate buffer also contained 80 mM 1,3-DMBA substrate. Plates were allowed to incubate for 30 min at room temperature before absorbance readings.

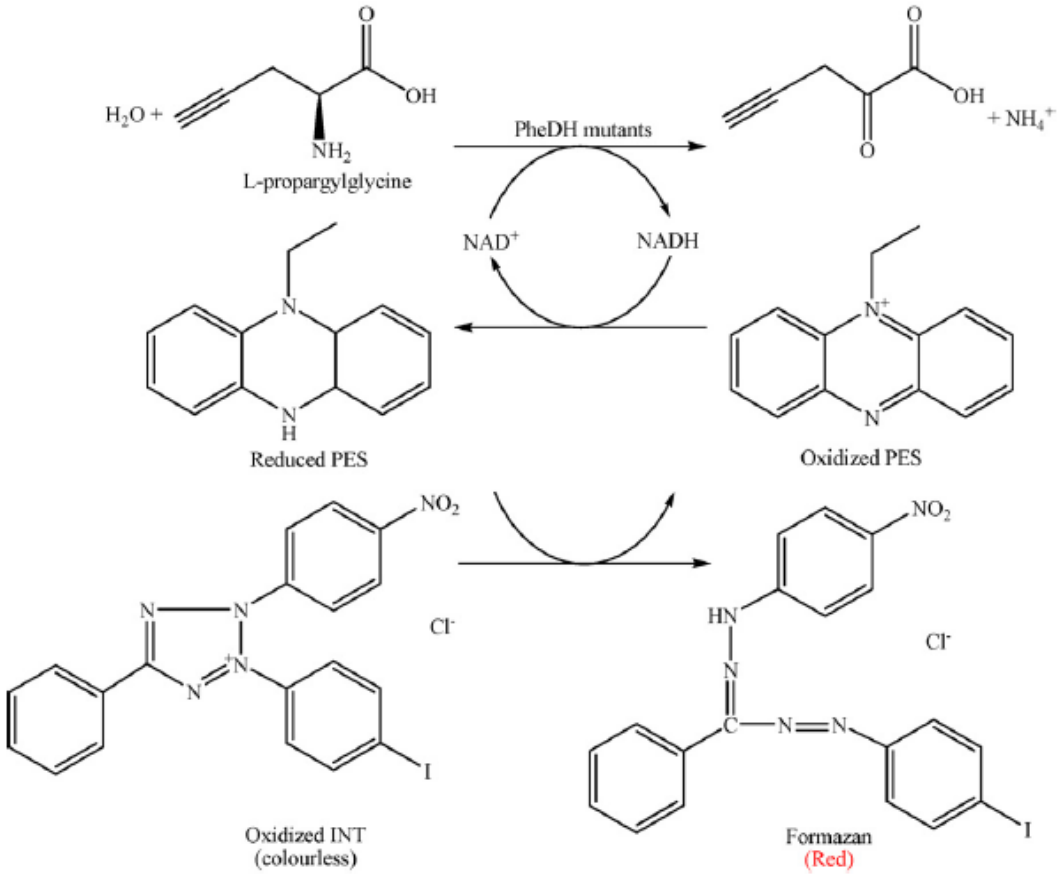


Figure 3.8. Reaction scheme of formazan colorimetric assay (88).

The final and most simple assay (Figure 3.9) was used in the analysis of libraries 8 through 11. The assay involved reading the wells' absorbences at two wavelengths, 340 nm and 600 nm. The increased absorbance at 340 nm corresponds to the production of NADH in the deamination of 1,3-DMBA, while the 600 nm reading roughly estimates

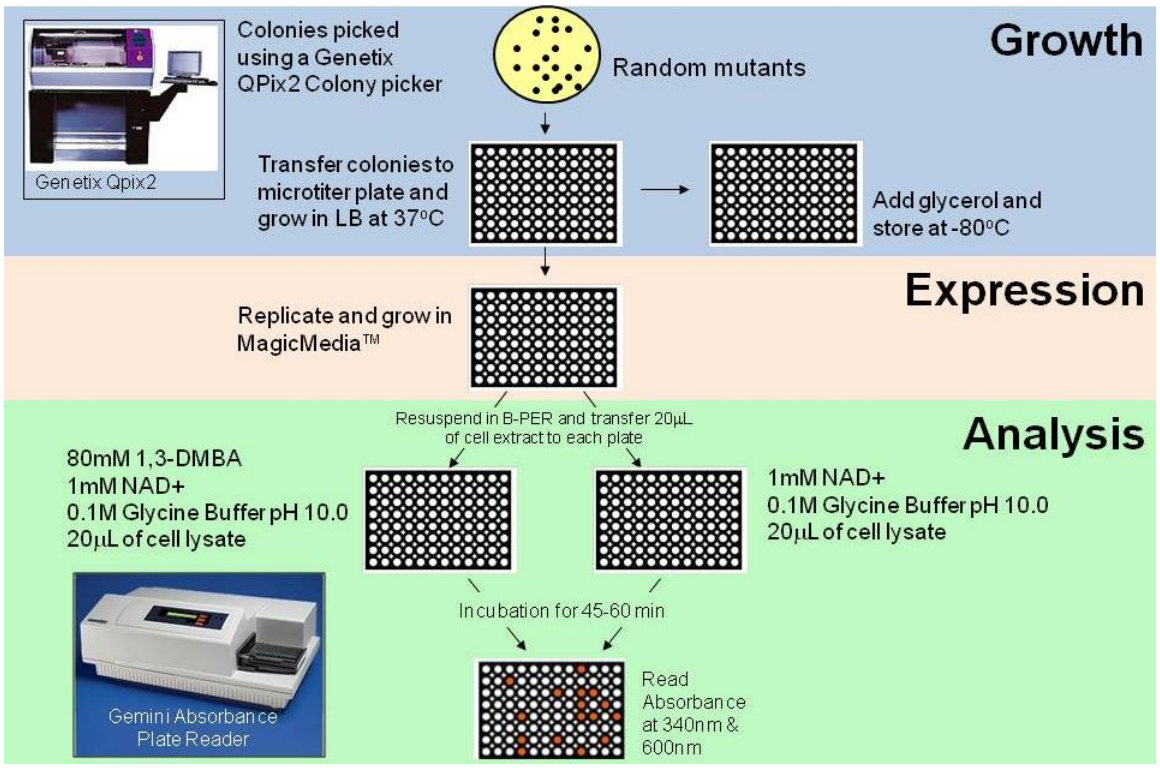


Figure 3.9. 340 nm absorbance-based amine detection assay with correlation of background data to 600 nm for screening libraries 8 through 11.

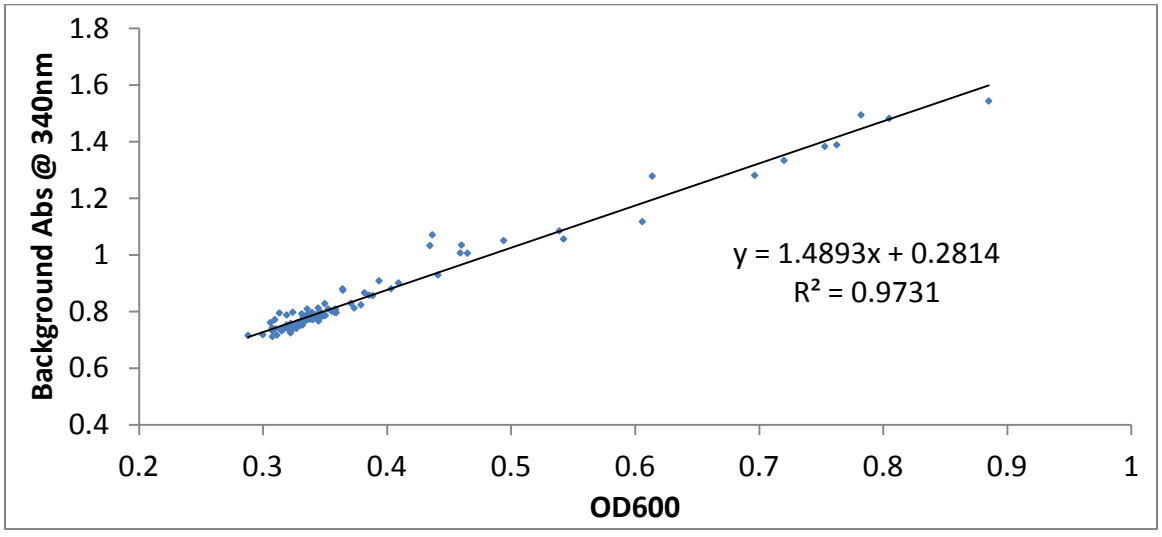


Figure 3.10. Correlation of LeuDH variants' OD600 absorbance to background absorbance of 340 nm.

the biomass present in the well. The absorbance at 600 nm correlates strongly the background absorbance at 340 nm. (Figure 3.10)

The absorbance at 340 nm was normalized by the 600 nm absorbance reading for both the reaction and background plates. Variants were then scored by their proportional increase over the background observed in each well (Equation 3.1). This method significantly reduces the background noise of the assay by accounting for differences in expression levels, causing more accurate determination of enzymatic activity and decreasing false positives. The resulting quantity roughly estimates the specific activity of the overexpressed mutant protein.

$$\frac{340\text{nm}\,R/B}{600\text{nm}\,R/B} = \frac{\text{Ratio of activity}\,\dfrac{\text{Reaction}}{\text{Background}}}{\text{Ratio of biomass}\,\dfrac{\text{Reaction}}{\text{Background}}} \qquad (3.1)$$

3.3 Results and Discussion

3.3.1 Identification of key residues in LeuDH

Leucine dehydrogenase from *Bacillus stereothermophilus* served as the initial protein scaffold. The only crystal structure currently available for leucine dehydrogenase (LeuDH) is an apo crystal structure from *Bacillus sphaericus* (PDB: 1LEH) (89). The determination of substrate–binding pocket interactions was instead based upon the holo crystal structures of an analogous protein, phenylalanine dehydrogenase (PheDH) from *Rhodococcus sp.* M4 (PDB: 1C1D and 1BW9) (90, 91). LeuDH and PheDH have a nearly identical secondary structure, with a backbone RMSD of only 0.234 Å, thus allowing for reasonable estimation of the interactions of the substrate with the binding

pocket. By using these structures in conjunction with knowledge of the reaction mechanism, amino acid residues surrounding the L-phenylalanine substrate binding pocket were identified (Figure 3.11). As previously discovered by Sekimoto et al., residues Lys80 and Asp115 are essential to the catalytic mechanism of the enzyme (92). These residues were excluded from our mutational libraries to conserve activity.

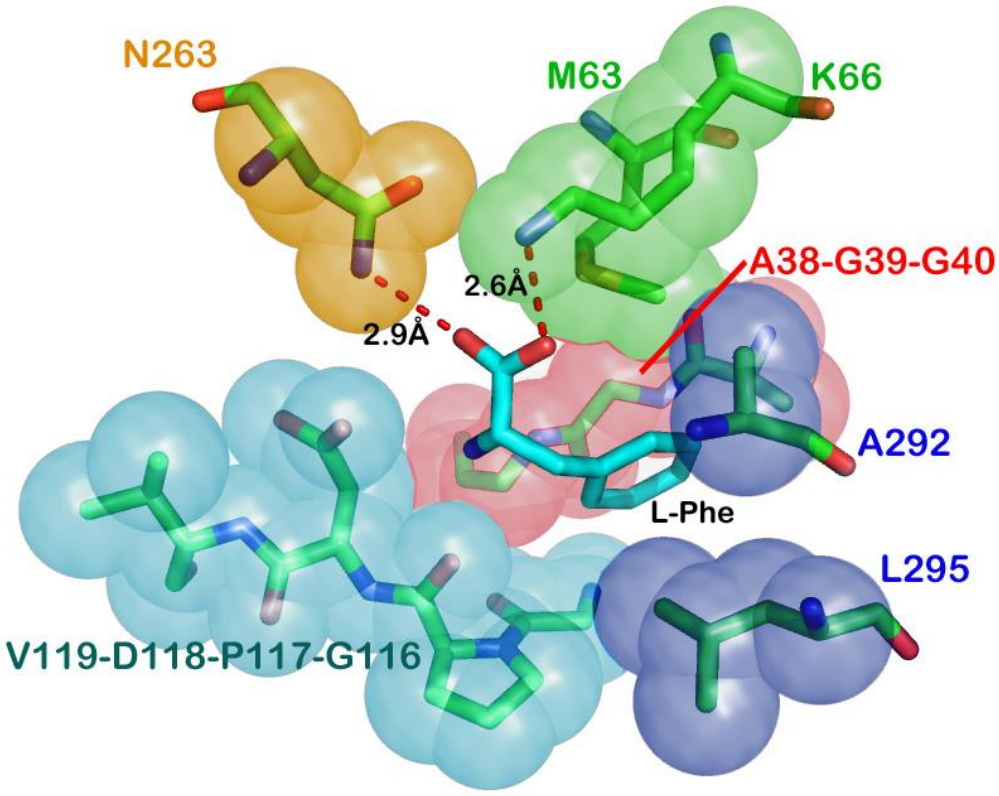


Figure 3.11. *Rhodococcus sp.* M4 PheDH active site with bound L-phenylalanine and surrounding residues suitable for mutation (PDB: 1C1D) (90).

Conversely, residue Lys67 of LeuDH (Figure 3.11, PheDH: K66) directly interacts with the carbonyl moiety of the amino acid substrate through favorable charge interactions, and was chosen as the initial point of mutation (90, 92). Library 1 included

mutation to each of the remaining 19 amino acids. The resulting variants were cloned into pET28a vector and BL21 (DE3) competent cells for subsequent expression and His-tag purification. Each purified variant was analyzed spectrophotometrically at 340 nm to correlate the NADH/NAD+ cofactor conversion with the amination and deamination activity (93).

Table 3.2. Grouping of active-site residues in *B. stearothermophilus* LeuDH. [a] Required screening for 95% statistical coverage of library sequence space. [b] Analyzed as individual mutants using purified protein.

Library	PheDH	LeuDH with degenerate codons	Library screening size[a]
1	K66	K67[b]	20
2	M63	M64 NNK	3066
	K66	K67 NNK	
3	G116	A112 GBC	3450
	P117	E113 NNT	
	D118	D114 RAB	
	V119	V115 RTK	
4	A38	L39 DBW	2910
	G39	G40 GBC	
	G40	G41 DBW	
5	S189	H187 NNK	94
6	A292	V290 DBS	969
	L295	V293 DBS	
7	G116	A112 NNK	94
8	I293	I291 NNK	94
9	E299	E296 NNK	94
10	N263	N261 NNK	94
11	K66	K67 DDK	969
	N263	N261 DDK	

The remaining binding pocket residues were broken into groups on the basis of CASTing principles to create libraries 2, 3, 4, and 6 (Table 3.2) (94). Regions of the binding pocket were grouped based upon their primary structure. The breadth of mutation at each position was then narrowed by their interactions based upon secondary structure. For example, every other amino acid side chain interacts with one another in a β-sheet, while only every fourth amino acid of an α-helix does the same. Interacting residues within each group were mutated simultaneously to capture any synergistic effects with neighboring residues. The mutational breadth at each position was constrained through degenerate codons to avoid amino acid substitutions which were likely detrimental to substrate binding (such as clear steric hindrance or unfavorable charge interactions). Removing a subset of detrimental amino acids limits the library size, so as to improve screening efficiency, while maintaining an increased chance of altering the substrate specificity. The number of colonies required to achieve 95% statistical coverage was calculated using equation 3.2 (31).

$$N = \frac{\ln(1-P)}{\ln\left(1-\frac{1}{n}\right)}; \qquad (3.2)$$

N = number of colonies required in screening,
P = statistical percent of coverage of library,
$\frac{1}{n}$ = proportion of the least represented codon

An example set of the degenerate codons applied in library 3 demonstrates the breadth of mutation over the four residues; Ala112, Glu113, Asp114, and Val115 (Table 3.3). These four residues create a loop region along the binding pocket. The side chains

of residues 113 and 115 face toward the binding pocket, while 112 and 114 are directed outward. Residues 113 and 115 are believed to have strong interactions in stabilizing the imine intermediate, and for this reason will be widely mutated excluding only mutations which create unfavorable charge interactions or drastic steric hindrance. Conversely, residues 113 and 115 are likely to have less interaction. These residues' mutations will be restricted to amino acids similar to the wild type residues, limiting mutational breadth to greatly reduce the required screening effort.

Table 3.3. LeuDH library 3 distribution of degenerate codon amino acids at positions 113 through 116.

Residue Codon	**A112 GBG**	**E113 NNT**	**D114 RAB**	**V115 RTK**
Ala [A]	**1**	**1**	0	0
Arg [R]	0	**1**	0	0
Asn [N]	0	**1**	**2**	0
Asp [D]	0	**1**	**2**	0
Cys [C]	0	**1**	0	0
Gln [Q]	0	0	0	0
Glu [E]	0	0	**1**	0
Gly [G]	**1**	**1**	0	0
His [H]	0	**1**	0	0
Ile [I]	0	**1**	0	**1**
Leu [L]	0	**1**	0	0
Lys [K]	0	0	**1**	0
Met [M]	0	0	0	**1**
Phe [F]	0	**1**	0	0
Pro [P]	0	**1**	0	0
Ser [S]	0	**2**	0	0
Thr [T]	0	**1**	0	0
Trp [W]	0	0	0	0
Tyr [Y]	0	**1**	0	0
Val [V]	**1**	**1**	0	**2**
Stop	0	0	0	0
Codons	**3**	**16**	**6**	**4**
AA	**3**	**15**	**4**	**3**

Single residue libraries (5, 7, 8, and 9) were broadly mutated to all 20 amino acids (codon: NNK), because of their small size and the insignificant decrease in the required screening effort by applying a more restrictive degenerate codon. These residues were previously identified as influential in changing the substrate specificity of *Bacillus sphaericus* PheDH (88, 95).

Library 10 included mutation of binding pocket residue N261 (Figure 3.11, PheDH: N263). This position was not initially considered for mutation, despite its interaction with the carboxy moiety of the wild-type ligand, because of its involvement in binding the cofactor. However, as only the backbone amino group of the residue interacts with the cofactor, an exception was made for mutation at this position (Figure 3.12). The desired protein interactions with the cofactor will be mostly preserved after substitution of the side chain due to the universal peptide backbone structure. Being a single residue library, N261 was broadly mutated to all 20 amino acids (codon: NNK)

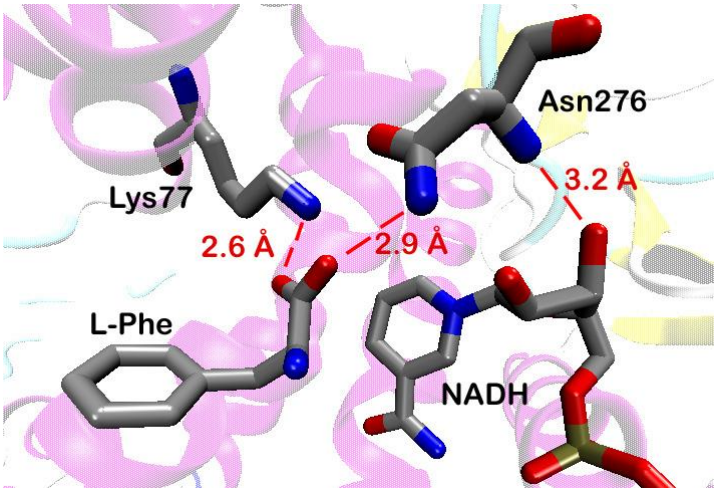


Figure 3.12. Asn276 side chain interactions with the L-Phe substrate and backbone interactions with NADH.

Lastly, residues K67 and N261 were mutated simultaneously to create library 11 because of their proximity in the binding pocket to the carboxyl moiety of the wild-type keto-acid substrate (2.6 Å and 2.9 Å, respectively). These residues have a synergistic effect with respect to binding the wild-type substrate's carboxyl moiety, and interact directly with one another.

3.3.2 Screening results via the high-throughput assay

Prior to screening, each library was checked for diversity at the target residues by sequencing randomly selected colonies (10-20 colonies), and comparing their distribution to the applied degenerate codon. Table 3.4 gives an example of 20 randomly selected wells from library. The diversity at positions 39-41 indicates successful application of the mutagenesis protocol. Similar diversity was observed for all libraries.

The NAD+ auto-fluorescence assay (libraries 2-5) typically indicated a net negative value for non-beneficial mutations; the net relative fluorescence units (RFU) ranged from 0 to -500. Beneficial mutations gave net positive RFU, and were easily identifiable over background.

Figure 3.13 contains an example of a typical well plate analyzed using the absorbance-based assay with readings at 340 nm and 600 nm. Normalized absorbance values were calculated using Equation 3.1. Unsuccessful mutants show a normalized absorbance of approximately unity indicating no enhancement over background, while the successful mutants indicated the desired catalytic activity with normalized absorbance beyond 1.5.

Table 3.4. Sequencing results of random LeuDH library 4 colonies to confirm diversity after the application of degenerate codons.

Residue	39	40	41
WT	**Leu**	**Gly**	**Gly**
1	Ala	Val	Ala
2	STP	Val	Ser
3	Phe	Ala	Ile
4	Ala	Val	Ala
5	Val	Gly	Thr
6	Ala	Val	Ser
7	Arg	Ala	Ser
8	Val	Gly	Thr
9	Cys	Gly	Thr
10	Arg	Val	Ser
11	Thr	Gly	Ile
12	Thr	Val	Val
13	Cys	Val	Ile
14	Val	Gly	Ser
15	Ala	Arg	Thr
16	Thr	Leu	Ser
17	Ile	Gly	STP
18	Gly	Gly	Phe
19	Val	Ala	Val
20	Ala	Gly	Ser

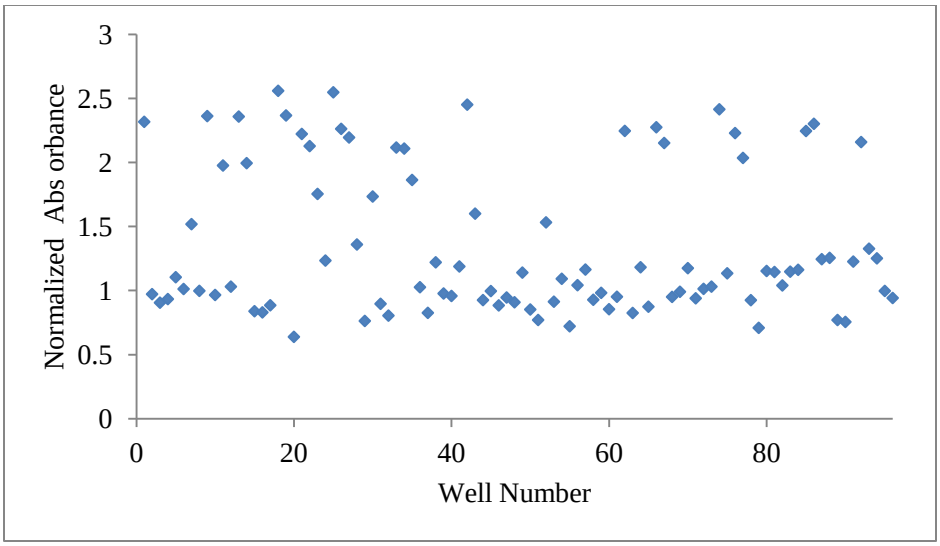


Figure 3.13. LeuDH library 11, plate 9. An example of the results obtained for the high-throughput absorbance-based assay.

3.3.3 Characterization of library hits

The individual analysis of library 1 variants yielded a single beneficial mutation, Lys67Met that exhibited low, yet unprecedented activity for reductive amination (0.2 mU mg^{-1}). A higher activity of 3.4 mU mg^{-1} was observed in the oxidative deamination of racemic 1,3-dimethyl butyl amine (1,3- DMBA). This mutation was carried over into the subsequent libraries.

The top performers of each high-throughput library (libraries 2-11) were sent for sequencing. The resulting sequences resulted in either the identification of beneficial mutations, or a large presence of parental variants. When beneficial mutations were identified, they often appeared multiple times in the sequenced hits. (Table 3.5) The assay accurately and repeatedly identified beneficial mutations. When all or mostly parental variants were identified as the best performing mutants, these libraries failed to

Table 3.5. LeuDH library 3 hit sequencing. These sequences represent the top 10 performing colonies as indicated by the high-throughput assay. Repeat identifications of beneficial mutations are grouped at the top of the table. The top variants also returned the wild type sequence (*italics*).

Residue			
A112	**E113**	**D114**	**V115**
Ala	Val	Asp	Val
Ala	Val	Asp	Val
Ala	Val	Asp	Val
Ala	Leu	Asp	Val
Ala	Leu	Asp	Val
Val	Ala	Asp	Met
Gly	Phe	Asn	Ile
Ala	Phe	Asn	Val
Ala	*Glu*	*Asp*	*Val*
Ala	Cys	Asp	Val

identify further mutations. The enrichment of parental variants from a random library further supports the accuracy and robustness of the successful high-throughput screen.

To get a more detailed determination of activity, each hit was further evaluated as purified protein. The hits were expressed and purified as described in Section 3.2.2. The purified protein's activity was measured over a range of MIBK and 1,3-DMBA concentrations to determine k_{cat} and K_M of the amination and deamination reactions, respectively (Section 3.2.3). An example of these results can be seen in Table 3.6 for the analysis of library 11.

Table 3.6. Characterization of LeuDH library 11 hits and comparison to previous best preforming variants from library 10. n.m.= not measurable (< 0.1 mU mg^{-1})

	Residue		Amination of MIBK		Deamination of 1,3-DMBA	
Well	**K67**	**N261**	**V_{max}**	**K_M**	**V_{max}**	**K_M**
Previous Top Variants						
Lib10 B2	Met	Cys	0.236 ±0.04	21.6 ±11.5	1.76 ±0.20	14.8 ±6.1
Lib10 H3	Met	Val	0.084 ±0.01	21.01 ±8.5	**2.894 ±0.44**	**33.1 ±13.0**
Library 11 Hits						
2B1	Ser	Leu	**0.331 ±0.07**	**18.7 ±11.7**	2.350 ±0.26	41.9 ±9.7
1C9	Ser	Met	0.326 ±0.02	61.1 ±6.2	1.381 ±0.17	28.3 ±8.3
1H4	Val	Met	0.240 ±0.04	37.6 ±14.4	4.841 ±3.44	251.7 ±224
2B11	Gly	Leu	0.211 ±0.07	37.7 ±27.2	1.710 ±0.17	111.4 ±17.7
1G2	Leu	Leu	0.159 ±0.04	27.2 ±19.0	6.688 ±18	498.0 ±1551
1A2	Leu	Val	0.022 ±0.00	16.7 ±8.0	0.976	33.8
2E12	Gly	Gly	2.4E+09	3.8E+13	n.m.	n.m.
2G8	Gly	Cys	7.2E+15	8.2E+18	5.1E+11	7.0E+13

Table 3.7. Accumulated mutations and resulting improvements to the specific activity and K_M of LeuDH. Error values represent 95% confidence intervals of nonlinear fit parameters. [a] Maximum specific activity (U mg^{-1} protein). [b] MIBK or 1,3-DMBA substrate (mM). [c] n.m.=not measurable (< 0.1 mU mg^{-1}). [d] Separate variants gave maximum amination and deamination activity.

Library	Mutations	Reductive amination specific activity[a]	K_M[b]	Oxidative deamination specific activity[a]	K_M[b]
WT	-	n.m.[c]	n.m.	n.m.	n.m.
1	K67M	0.0002	n.m.	0.0034	n.m.
3	K67M, E113V	0.015	n.m.	0.65	48
6	K67M, E113V, V290C	0.016 ±0.004	70.2 ±35	1.02 ±0.41	77.1 ±66
10[d]	K67M, E113V, N261V, V290C	0.089 ±0.007	10.3 ±3.6	**2.81 ±0.44**	**30.9 ±13**
	K67M, E113V, N261C, V290C	0.236 ±0.04	21.6 ±12	1.76 ±0.20	14.8 ±6.0
11	K67S, E113V, N261L, V290C	**0.690 ±0.07**	**15.1 ±5.1**	2.64 ±0.28	57.5 ±12.5

After the screening of 11 libraries, four beneficial mutations were identified. Early rounds of mutation (libraries 1–9) identified either the wild-type sequence or single variants which improved the activity simultaneously in both the amination and deamination directions, thus allowing for the straightforward selection of the top variant. Library 10 identified distinct mutations for the most-active amination and deamination variants (Table 3.7). The most active amination and deamination mutations were with position Asn261 exchanged by Cys and Val, respectively. Library 11 further improved the amination activity by identifying synergistic mutations at positions 67 and 261, with these positions mutated from those identified in libraries 1 and 10. Libraries 2, 4, 5, and 7-9 were unsuccessful in indentifying beneficial mutations, returning the parental sequence as hits. The final amine dehydrogenase contained four mutations, which are summarized in Table 3.7. The resulting quadruple variant (K67S/E113V/N261L/V290C) gave the highest amination activity of 0.69 U mg^{-1} with a corresponding k_{cat} value of 0.46 s^{-1}, as determined by a non-linear fit (Figure 3.14) of the Michaelis-Menten equation (3.3) using MATLAB;

$$v = \frac{d[P]}{dt} = \frac{V_{max}[S]}{K_M + [S]}; \qquad (3.3)$$

v = reaction rate,
[P] = product concentration,
[S] = substrate concentration,
V_{max} = maximum rate under saturated conditions

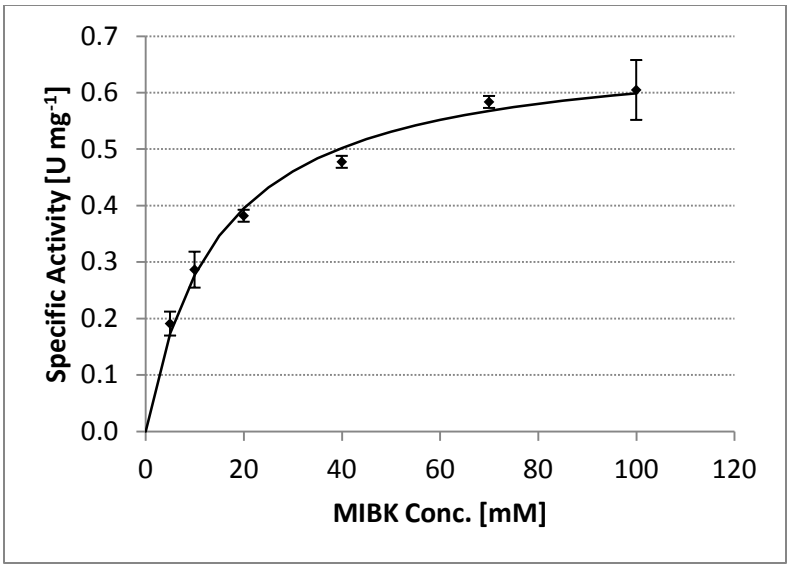


Figure 3.14. Non-linear Michaelis-Menten fit of MIBK amination activity by LeuDH-AmDH quadruple variant, K67S/E113V/N261L/V290C.

3.4 Detailed characterization of library 11 hit

3.4.1 Substrate specificity

The most-active variants showed amination and deamination activity toward a number of ketones and amines, respectively (Table 3.8). Since MIBK most closely mimics the wild-type keto-acid, 3-oxo-2-pentanoate, it was not surprisingly the most activity of the ketones tested. The detailed measurements of the final variant (K67S/E113V/N261L/V290C) kinetics showed enhanced amination activity over the previous library 10 mutant toward all the ketones investigated, except methyl acetoacetate.

Table 3.8. Substrate profile of top amination and deamination variants. [a]Activity measured with 20 mM substrate. [b]Specific activity in mU mg^{-1} protein, with the error representing one standard deviation. [c]Activity measured with 40 mM racemic MBA, 20 mM of each enantiomer.

Substrate[a]	K67M/E113V/N261V/V290C[b]	K67S/E113V/N261L/V290C[b]
(R)-MBA	476.5 ±1.4	586.3 ±4.1
(S)-MBA	5.0 ±0.0	1.6 ±0.0
(R/S)-MBA[c]	484.0 ±3.5	784.6 ±13.4
cyclohexylamine	-	56.0 ±0.0
cyclohexanone	18.8 ±0.0	123.4 ±3.5
ethyl pyruvate	19.8 ±7.0	13.2 ±5.8
methyl acetoacetate	4.5 ±0.7	4.5 ±0.6
ethyl-3-oxohexanoate	6.4 ±4.9	14.0 ±1.2
acetophenone	3.5 ±0.7	58.8 ±1.7

3.4.2 Conversion and GDH cofactor recycle system

Initial rate kinetic measurements are used primarily in the determination of kinetic parameters to avoid second-order effects and product inhibition. The enzyme's ability to reach complete substrate conversion cannot be determined from these initial rate experiments, and was instead evaluated by chiral GC of overnight conversion with a GDH recycle system. The GDH regenerates the expensive NADH by redox reaction of NAD+ and inexpensive glucose (Figure 3.1). In 1.5 mL reactions, replicates containing 10 mM MIBK substrate was mixed with an excess of 12 mM glucose, 10 U GDH, and a catalytic amount of NADH (200 μM). Histag purified K67S/E113V/N261L/V290C LeuDH-AmDH (0.13 mg) was added to each reaction and allowed to convert for 18 hours. An additional 0.13 mg of enzyme was then added and allowed to react for 10 hours to achieve the maximum conversion. One mL aliquots of the final reaction media

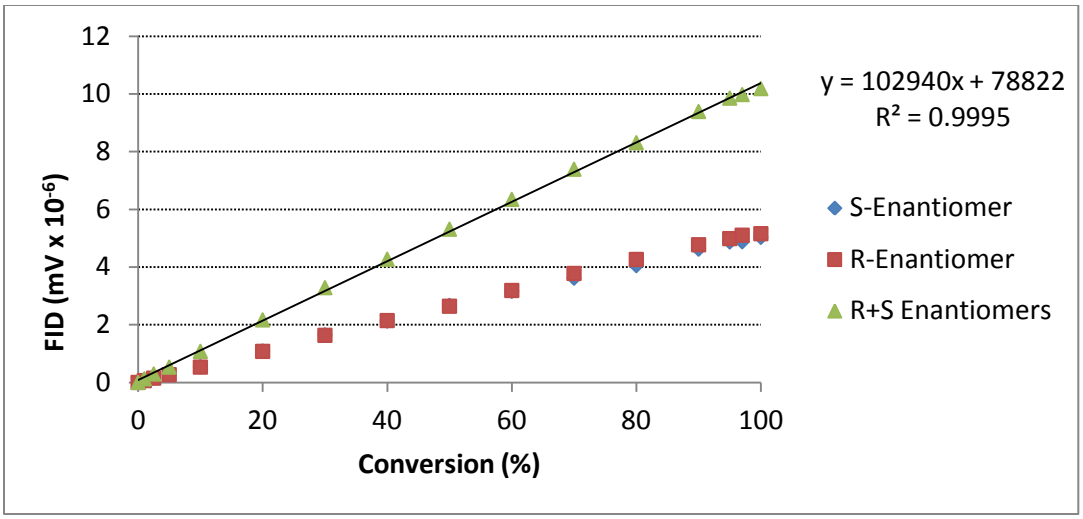


Figure 3.15. Conversion calibration curve of MIBK to 1,3-DMBA by LeuDH-AmDH.

were adjusted to pH >13 by the addition of 100 µL of 10 N NaOH and derivatized in toluene according to the TFAA protocol (96). The replicates were then compared against a calibration curve derived by the same procedure with simulated levels of conversion (Figure 3.15), and an average substrate conversion of 92.5%±2.6 was achieved.

3.4.3 ^{1}H NMR product confirmation

The product amine was additionally confirmed through ^{1}H NMR. A 175 mL-scale amination reaction with GDH cofactor recycle was performed similar to previous conversion experiments. The reaction media was a 500mM NH_4Cl/NH_4OH buffer at pH 9.6 containing; 10 mM MIBK, 12 mM glucose, 1 mM NADH, 2 mg GDH (85 U/mg). The histag purified quadruple variant (K67S/E113V/N261L/V290C @ 1314 µg/mL) was added in two 10 mL aliquots, initially and after 24 hrs. After 48 hours, the pH was increased (>12) with the addition 25 mL 10 N NaOH. The deprotonated amine product was then extracted with 40 mL methyl tert-butyl ether. Following separation, the solvent was removed through rotovap distillation, and the purified product was analyzed by ^{1}H

NMR. The reaction showed 84.7% conversion, yielding 29.5 mg (*R*)-1,3-DMBA (>99.3% e.e.) with spectra matching the standard.

^{1}H NMR (400 MHz, $CDCl_3$) δ 2.94 (sex, 1H, *J* = 6.4 Hz), 1.65 (sep, 1H, 6.8 Hz), 1.24-1.02 (m, 2H), 1.04 (d, 3H, 6.4 Hz), 0.89 (d, 6H, 6.8 Hz).

3.4.4 Enantioselectivity

The enantioselectivity was initially estimated by measuring the deamination activity toward individual enantiomers of methylbenzylamine (MBA). MBA was used in place of 1,3-DMBA for selectivity experiments since it had a reasonable level of activity and individual enantiomers of MBA were commercially available (Sigma Aldrich). A preference towards (*R*)-MBA over the corresponding (*S*)-enantiomer was evident in the discrepancies between the deamination activities of 0.586 U mg^{-1} and 0.002 U mg^{-1}, respectively. These preliminary results were corroborated through direct measurement of the enantioselectivity of the MIBK amination in producing chiral 1,3-DMBA.

Derivatization with trifluoroacetic anhydride (TFAA) (96) and chiral gas chromatography allowed for the separation and quantification of individual enantiomers. The derivatized 1,3-DMBA enantiomers gave adequate baseline separation of the peaks, with elution times of 47.5 and 48.9 min for the (*S*)- and (*R*)-enantiomers, respectively. (Figure 3.16) Overnight reactions were pH adjusted to 13 with 10 N NaOH and extracted from the aqueous phase into benzene. The amine was derivatized and concentrated through evaporation of benzene. The product was composed of the single (*R*)-enantiomer of DMBA. The product amine was injected in increasing concentrations until the (*S*)-

enantiomer could be observed. The quadruple variant exhibited an *e.e.* of 99.8%. Pure enantiomers of 1,3-DMBA are not available commerically. Therefore, the product enantiomer was identified by optical rotation. Optical rotations of each enantiomer are represented in Table 3.9. The extracted product amine in methanol gave an optical rotation of -0.0039º, indicating the production of the *(R)*-enantiomer (97). This level of enantioselectivity was maintained even beyond 90% substrate conversion.

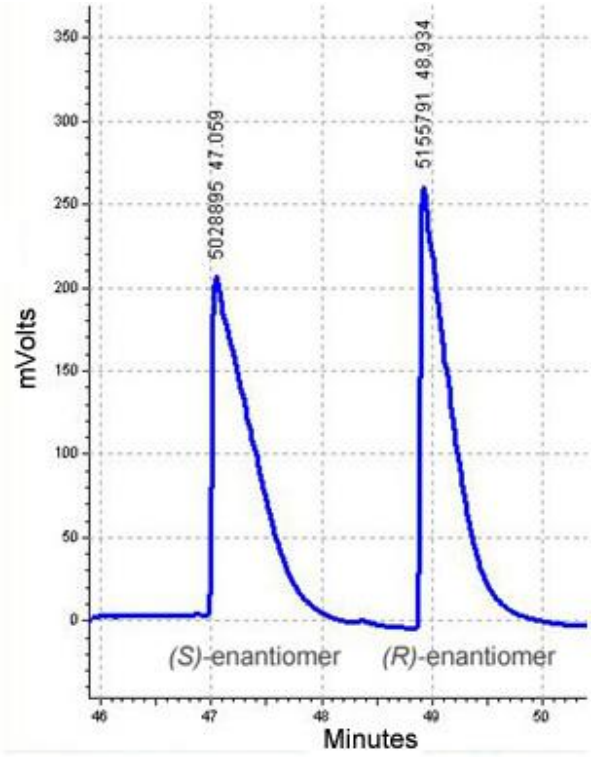


Figure 3.16. Chiral gas chromatography (FID) separation of derivatized racemic 1,3-DMBA standard.

Table 3.9. Optical rotations of 1,3-DMBA enantiomers, neat, at 589 nm (97).

Enantiomer of 1,3-DMBA	Optical rotation @ 589 nm (neat)
(R)-enantiomer	-11.2º
(S)-enantiomer	+11.2º

3.4.5 Thermostability

The thermostability of the top LeuDH-AmDH variants was compared to the wild type LeuDH by circular dichroism. Each protein was histag purified and dialyzed in 50 mM sodium phosphate buffer pH 8.0, 300 mM NaCl, removing the imidazole from the purification procedure. Each protein was diluted to 50 μg mL^{-1} in a 0.1 cm quartz cuvette for analysis. Ellipticity was measured at 222 nm, corresponding to the maximum ellipticity of the protein as indicated by the wavelength scan (Figure 3.17). Each protein was analyzed from 25 °C to 95 °C with a ramp rate of 1 °C min^{-1}. All three proteins exhibited nearly identical unfolding with melting temperatures (T_M) between 65 °C and 68 °C, making any differences in stability difficult to characterize (Figure 3.18). The four mutations within the LeuDH binding pocket did not destabilize the protein scaffold compared to the wild type.

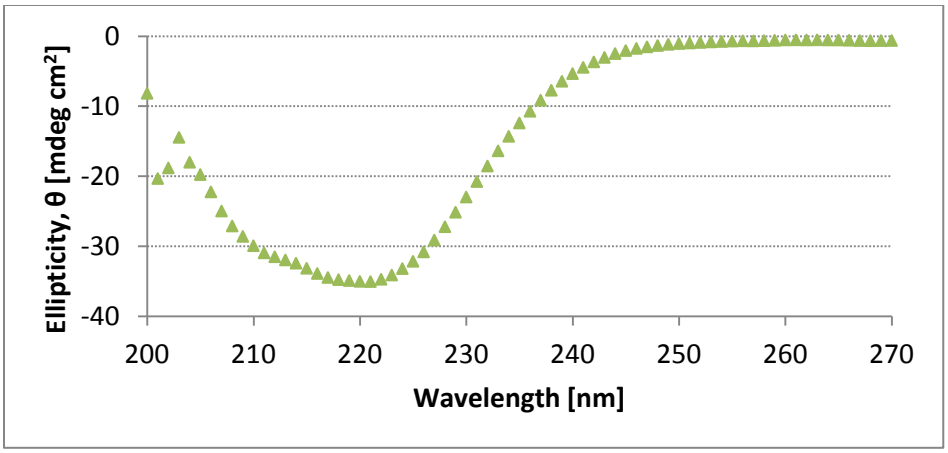


Figure 3.17. Circular dichroism wavelength scans of wild type LeuDH.

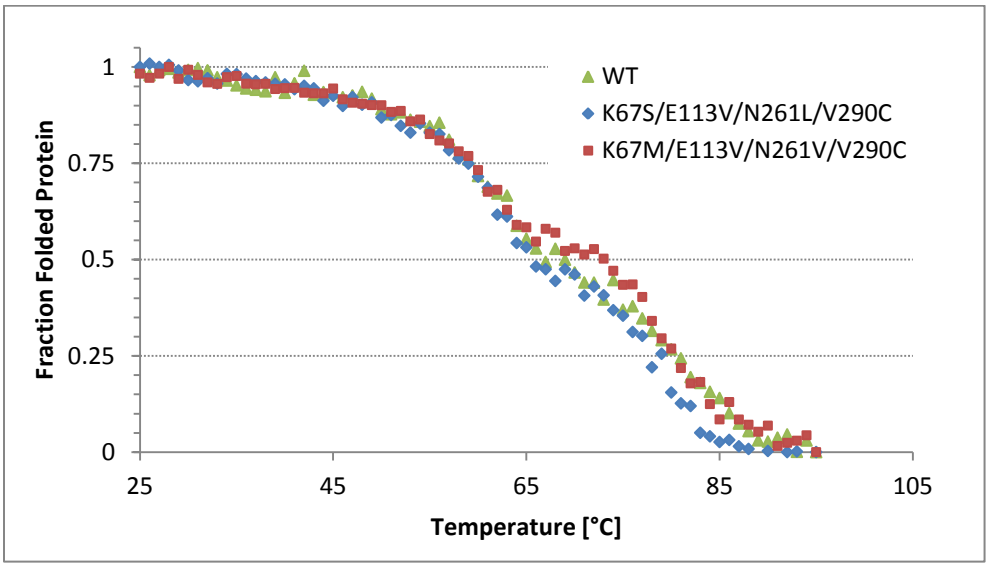


Figure 3.18. Circular dichroism of wild type LeuDH and top LeuDH-AmDH variants with normalized ellipticity, representing the percentage of folded protein.

3.5 Conclusions

An amine dehydrogenase was successfully developed from an existing amino acid dehydrogenase template by active-site targeted protein engineering. Eleven rounds of directed evolution have completely altered the enzyme's specificity and created amination activity. Each round of mutagenesis focused on a region of the binding pocket,

including simultaneous mutation of neighboring residues to capture synergistic effects. These variants were screened by various high-throughput assays identifying minute increases in amination activity, and successful mutations were carried into future rounds of mutagenesis.

The largest improvements in activity were achieved by the cooperative mutation of residues K67 and N261. Their simultaneous mutation ultimately indentified the most active quadruple mutant, K67S/E113V/N261L/V290C with novel reductive amination activity of 0.69 U mg^{-1} with a corresponding k_{cat} value of 0.46 s^{-1}. Within the first round of mutation, the native activity toward L-Leu of 112 U mg^{-1} (98) was greatly decreased to less than 2 mU mg^{-1}; completely inverting the enzyme's specificity. The enantioselectivity of the wild-type enzyme was maintained despite the drastic changes to the binding pocket and yielded (*R*)-1,3-DMBA with an *e.e.* value of 99.8% at 92.5% conversion. This amine dehydrogenase exhibited activity toward a number of different substrates. The enzyme has also maintained its wild type stability making it an attractive catalyst in the synthesis of chiral amines. This is the first example of a cofactor-dependent amine dehydrogenase capable of selectively synthesizing chiral amines from a prochiral ketone and free ammonia.

CHAPTER 4: DEVELOPMENT OF AN AMINE DEHYDROGENASE FROM PHENYLALANINE DEHYDROGENASE

4.1 Introduction

As the field of protein engineering grows, the characterization of novel biocatalysts expands the repertoire of allowable reactions. Biocatalysis is emerging as an essential tool in the asymmetric catalysis of pertinent chemicals. Chiral intermediates play a particularly important role as building blocks to the pharmaceutical industry. Enantiomerically pure forms of active pharmaceutical ingredients (APIs) can lead to lower dosages, increased efficacy, and elimination of detrimental side-effects (99). In 2006, 80% of small-molecule drugs approved by FDA were chiral and 75% were single enantiomers (2). The production of chiral amines is of particular importance due to their highly active and diverse influence on biological functions. Amine-based pharmaceuticals have been used in a wide range of functions including; stimulants, decongestants, vasoconstrictors, and antidepressants.

The production of chiral amines remains difficult through traditional chemical catalysis, and the direct amination of ketones with free ammonia to product chiral amines has been identified as one of the most aspired biocatalytic reactions challenging the pharmaceutical industry by the American Chemical Society's Green Chemistry Institute, Pharmaceutical Roundtable (9). Recently, this reaction has been demonstrated by an amine dehydrogenase (AmDH) through the modification of leucine dehydrogenase (LeuDH) scaffold to accept ketone substrates (79). This biocatalytic route of production

offers several advantages; elimination of heavy metals, high selectivity (99.8% *e.e.*) and conversion (> 92%), and improved atom economy.

Recently, alternative methods to the biocatalytic production of chiral amines have been shown to be effective. ω-Transaminases have been evolved to accept a number of ketone compounds, and are even being applied commercially in the production of sitagliptin (47, 63, 70, 71, 75, 76). ω-Transaminases potentially offer many of the same advantages as an AmDH with the exception of i) accepting free ammonia for amination, instead requiring the use of a sacrificial amine donor, and ii) overcoming equilibrium limitations to achieve full conversion.

In this chapter, we design a second example of an amine dehydrogenase catalyst based upon the scaffold of phenylalanine dehydrogenase (PheDH, EC 1.4.1.20). This PheDH-based amine dehydrogenase (PheDH-AmDH) offers several advantages including; increased specific activity over the LeuDH-based amine dehydrogenase (LeuDH-AmDH), amination of additional ketones, and increased tolerance to organic solvents.

4.2 Comparison of dehydrogenases and gene selection

Leucine dehydrogenase from *Bacillus stearothermophilus*, the enzyme detailed in Chapter 3, belongs to the Glu-Leu-Phe-Val dehydrogenase sub-family (Accession number: cd05211; Appendix A.1) of proteins. The enzymes of this family are closely related in structure and function. Each enzyme similarly catalyzes the reversible amination and deamination of their respective keto- and amino acids with free ammonia, and is facilitated by the hydride transfer of an NAD(P)H cofactor. These enzymes consist of two domains separated by a deep cleft (Figure 4.1). One domain is responsible of

binding the substrate and contains residues involved in catalysis (100, 101). The other proteins' domain responsible for binding the cofactor containing a characteristic βαβ structure, known as the Rossmann fold, and is highly conserved among these dehydrogenases (102). This conserved structure and functionality

makes the Glu-Leu-Phe-Val dehydrogenase sub-family part of the larger domain superfamily NADB_Rossmann (Accession number: cl09931); easily identifiable through the NCBI's Conserved Domain Database (103).

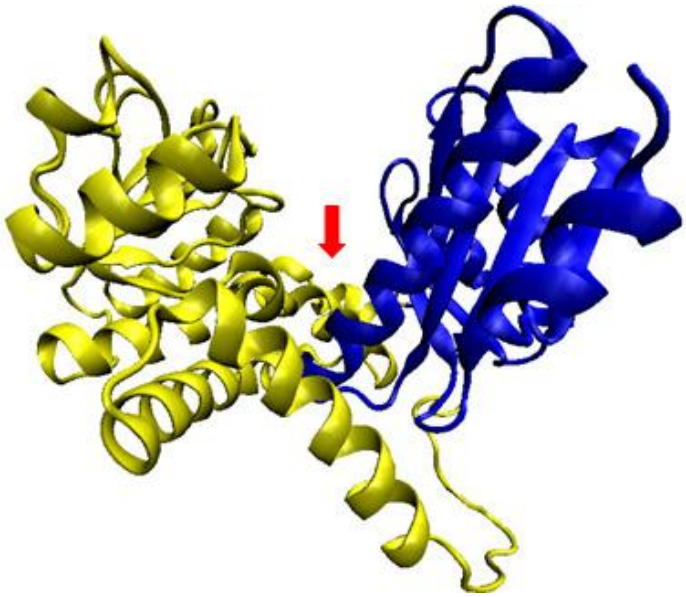


Figure 4.1. Crystal structure of LeuDH *Lysinibacillus sphaericus* (PDB:1LEH) (89). The catalytic region (blue) and Rossmann-fold region (yellow) are divided by a deep cleft. The active site lies within this cleft indicated by the red arrow.

A subset of phenylalanine dehydrogenases were selected from the Glu-Leu-Phe-Val dehydrogenase sub-family for comparison for use as a template for the PheDH-AmDH. The four organisms; *Bacillus badius* Bachelor, *Bacillus sphaericus*, *Rhodococcus sp.* M4, and *Sporosarcina ureae*, were initially selected because of their

well characterized activities and substrate profiles present in literature (90, 104-108). The similarity and identity between these genes can be seen in Table 4.1. Ultimately, *Rhodococcus sp*. M4 was chosen because of the availability of several crystal structures with bound substrates (PDB: 1C1D, 1C1X, 1BW9, and 1BXG) (90, 91), and its high specific activity (100). This allows for detailed interactions of the enzyme-substrate complex to be drawn directly from the crystal structures.

Table 4.1. Similarity and identity of select PheDHs with *Bacillus stearothermophilus* LeuDH. [% similarity/% identity]

Similarity/Identity	LeuDH *Bacillus stearothermophilus*	PheDH *Rhodococcus sp.* M4	PheDH *Bacillus badius*	PheDH *Bacillus sphaericus*	PheDH *Sporosarcina ureae*
LeuDH *Bacillus stearothermophilus*	-	54/36	66/48	71/50	69/48
PheDH *Rhodococcus sp.* M4		-	49/32	49/32	51/31
PheDH *Bacillus badius*			-	83/67	82/66
PheDH *Bacillus sphaericus*				-	78/63
PheDH *Sporosarcina ureae*					-

4.3 Evolution of phenylalanine dehydrogenase from *Rhodococcus sp.* M4

4.3.1 Wild-type gene synthesis, cloning, and overexpression

The wild-type gene sequence of phenylalanine dehydrogenase from *Rhodococcus sp*. M4 was attained from Brunhuber et al. (100). The nucleic acid sequence was synthesized by an external vendor (MWG Operon) and provided in a pUC57 vector. The

gene was amplified out of the host plasmid, along with the simultaneous insertion of restriction sites *Nde*I and *Hind*III (bold) using the primers; N-terminal 5'-GGGAATTC**CATATG**AGTATCGACAGCGCACTGAAC -3', and C-terminal 5'-CCATGATTACGCC**AAGCTT**GC -3'. A second amplification was performed to insert a STOP codon at the C-terminus of the protein sequence and allow for proper translation and expression. The forward and reverse primers, 5'-CGACAACGACAGCGACTGCCTAGGGCCCGTCGACTGCAG -3' and 5'-CTGCAGTCGACGGGCCCTAGGCAGTCGCTGTCGTTGTCG -3' respectively, were used in this amplification. The final insert was then digested and ligated into pET28a for overexpression and histag purification as described in Sections 3.2.1 and 3.2.2. The wild-type protein was successfully purified and activity confirmed (39.5 kDa, 46 U mg^{-1}).

4.3.2 Mutation and overexpression of AmDH

The wild-type gene was then mutated with the analogous mutations beneficial to creating AmDH in the LeuDH scaffold. These residues were identified through sequence alignment as shown in Figure 4.2.

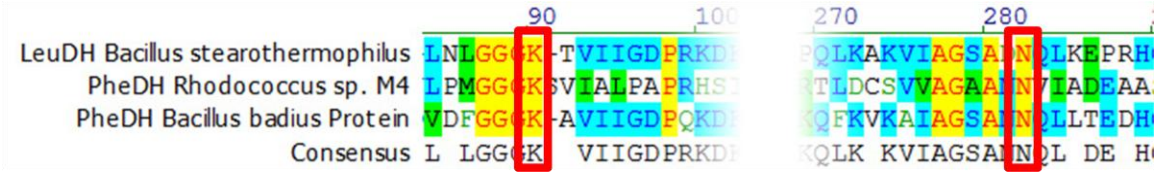


Figure 4.2. Truncated sequence alignment of LeuDH *B. stearothermophilus* and select PheDH showing consensus at pertinent residues LeuDH K67 and N261.

The K66M mutation was created by applying the primers, 5'-GGGGGCGATGACGTTG**ATG**ATGGCAGTGAGCAACCTTCC -3' and 5'-GGAAGGTTGCTCACTGCCAT**CAT**CAACGTCATCGCCCCC -3', in the overlap extension protocol described in Sambrook *et al.* (83). After sequence confirmation, overexpression of the K66M single mutant was attempted in pET28a vector and BL21 competent cells. This overexpression was unsuccessful, and the protein was not expressed. The K66M mutant DNA was then used as a template for mutation back to the parental sequence, K66. Upon mutation back to the parental sequence, successful overexpression of the protein was restored. This indicated that the *Rhodococcus sp.* M4 PheDH was not tolerant to mutation at this position without introducing difficulties in expression. An alternative PheDH gene from *Bacillus badius* was used in future experiments to circumvent these complex problems with overexpression.

4.4 Evolution of phenylalanine dehydrogenase from *Bacillus badius* Bachelor

Phenylalanine dehydrogenase from *Bacillus badius* Bachelor was chosen as an alternative scaffold to the *Rhodococcus sp.* M4 gene. The *B. badius* gene has been shown to be reasonably active with a k_{cat} value of 39 s^{-1} and has been extensively characterized in literature (105, 109-111). Another advantage of the *B. badius* PheDH is its increased similarity to the LeuDH scaffold (66% similarity) compared to the *Rhodococcus sp.* M4 (54% similarity). This increased the chances of successful translation of AmDH mutations from the previous scaffold.

The *Rhodococcus sp.* M4 and *B. badius* PheDHs are very different from each other on a sequence level, sharing only 32% identity and 50% similarity. Coming from the same structural sub-family, they share a much stronger correlation in their folding

motifs. Since there currently is no crystal structure available for the *B. badius* PheDH, the sequence was folded over the 1BW9 crystal structure and superimposed over the original scaffold using SuperPose (112). The resulting *B. badius* structure is nearly identical to the 1BW9 crystal structure with a backbone RMSD of only 1.69 Å (Figure 4.3), with the major deviation occurring at a loop region evident in the primary sequence. This tertiary structure similarity attests to the use of the 1BW9 crystal structure as a model of protein-substrate interactions.

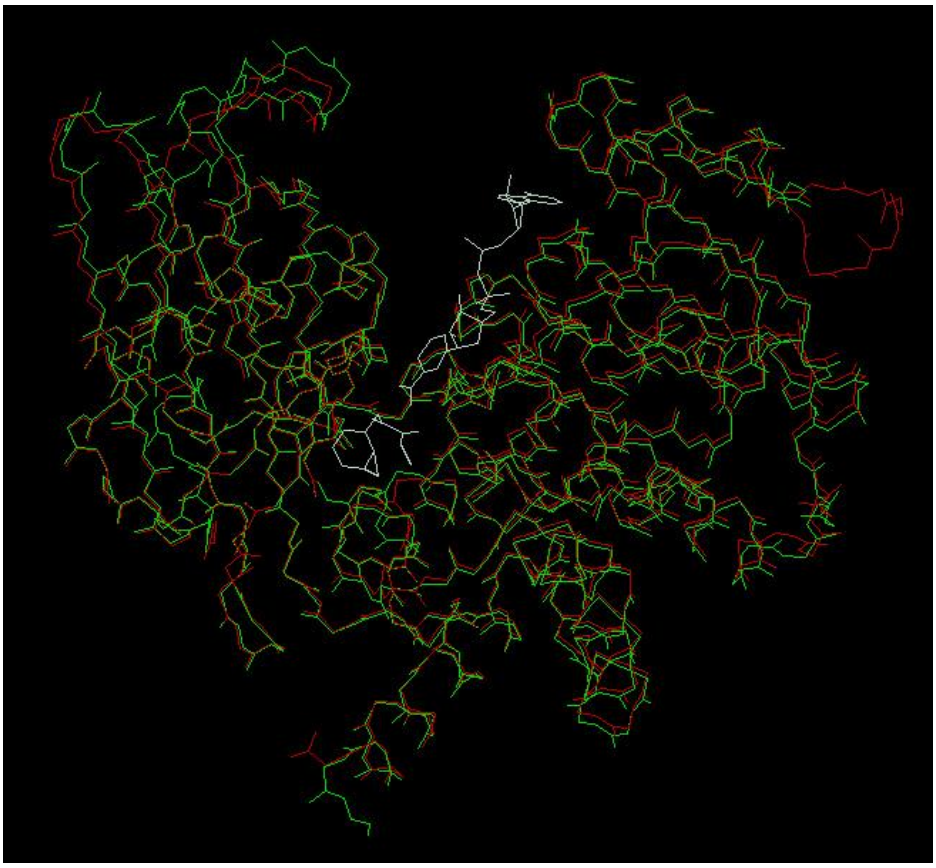


Figure 4.3. SuperPose alignment of *Rhodococcus sp.* M4 PheDH, green (PDB: 1BW9) and *B. badius* PheDH structural overlay, red. The NADH and L-Phe substrates are shown in white. The major deviation at the loop region is seen in the top right corner (112).

4.4.1 Materials and Methods

4.4.1.1 gDNA preparation, gene isolation and overexpression

Genomic DNA from *Bacillus badius* Bachelor (ATCC# 14574) was purchased from the American Type Culture Collection (ATCC, Manassas, VA). The PheDH gene could not be successfully amplified out of the fully intact genome. The genome had to instead be digested prior to amplification using restriction sites; *Bam*HI, *Bgl*II, *Eco*RI, and *Nde*I. These restriction sites are known not to cut within the PheDH gene, but would cut at several instances within the organism's genomic DNA. This allowed for a more manageable DNA template to amplify the PheDH gene. The gene was amplified using a standard PCR protocol (83) with the forward primer, 5'-GGAATTC**CATATG**AGCTTAGTAGAAAAAACATCCATCA -3' and reverse primer, 5'- CCG**CTCGAG**TATTAGTTGCGAATATCCCATTTTG -3'. These primers simultaneously inserted the restriction sites *Nde*I and *Xho*I prior to and following the gene, respectively. The locations of the restriction sites are indicated by bold font with each primer. This made for simple digestion and ligation into either pET17b or pET28a plasmids.

4.4.1.2 Protein purification

PheDH from *Bacillus badius* was expressed and purified in a manner analogous to that described in Section 3.2.2. Successful IMAC purification with the histag was confirmed through SDS-PAGE of the purification fractions (Figure 4.4). Overexpression of the enzyme is indicated by a heavy band at 44 kDa in the clarified cell lysate and protein fractions. The removal of cellular proteins can be seen in the wash fractions. Faint

bands of these proteins are observable in the purified fraction only after overloading. Under normal loading conditions, only the purified PheDH band is visible.

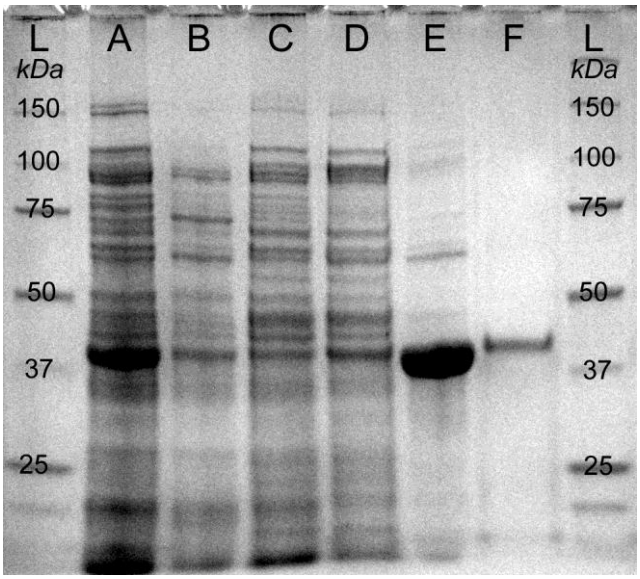


Figure 4.4. Histag purification of wild type *B. badius* PheDH. Lane L: Ladder, A: Clarified cell lysate, B: Unbound clarified cell lysate, C: 50 mM imidazole wash 1, D: 50 mM imidazole wash 2, E: Pure protein elution-overloaded, F: Pure protein elution-standard loading.

4.4.1.3 Spectrophotometric assay

The activity of the purified protein was evaluated using a spectrophotometric assay at 340 nm as described in Section 3.2.3 at 25 °C in various concentrations of NH_4Cl/NH_4OH buffer pH 9.6.

4.4.1.4 Mutagenesis and library generation

After eleven rounds of mutagenesis, three mutations were previously identified to be most influential in creating reductive amination activity in a LeuDH-based scaffold; Lys67Met, Glu113Val, and Asn261Val (Chapter 3). Analogous mutations to these were

identified through sequence alignment and applied to a similar amino acid dehydrogenase scaffold, phenylalanine dehydrogenase (PheDH) from *Bacillus badius* (48% identity, 66% similarity). Each of the point mutations, as well as combinations of these mutations, was evaluated for amination activity.

Mutants and libraries were generated using an analogous protocol to that described in Section 3.2.4, with the only difference that the restriction site *Xho*I was used in place of *Hind*III. Single and multiple variants containing mutations analogous to those found in 'LeuDH-AmDH library 10' were created using combinations of the primers in Table 4.2. LeuDH-AmDH library 11 had not yet been screened at the time of this experiment. The final variants were amplified from the respective mutational portions to yield the complete gene as described in the overlap extension protocol (83), and ligated into pET28a for subsequent expression and purification as detailed in Section 3.2.2.

The first library of PheDH-AmDH variants contained mutations at both positions K77 and N276. The directly substituted mutations from LeuDH-AmDH did not necessarily identify the optimal combination of mutations at these positions.

Table 4.2. Mutational primers applied to *B. badius* PheDH to create mutations analogous to those observed in LeuDH-AmDH.

LeuDH Mutant	PheDH Mutant		Primer	
K67M	K77M	Fwd	5' GCCTTTCCAAAGGAATGACTTAC**ATG**TGCGCGGCGTCCG	3'
		Rev	5' CGGACGCCGCGCA**CAT**GTAAGTCATTCCTTTGGAAAGGC	3'
E113V	T123V	Fwd	5' GGCGGCCGTTTCTATACAGGT**GTG**GATATGGGAACGAATATGGAAGATTTC	3'
		Rev	5' GAAATCTTCCATATTCGTTCCCATATC**CAC**ACCTGTATAGAAACGGCCGCC	3'
N261V	N276V	Fwd	5' GCAATCGCCGGTTCAGCC**GTG**AATCAGCTGCTTACGGAGGATCAC	3'
		Rev	5' GTGATCCTCCGTAAGCAGCTGATT**CAC**GGCTGAACCGGCGATTGC	3'

Simultaneous mutation of these residues will identify any synergistic effects between their side chains. Since previous libraries had identified LeuDH K67M as a beneficial mutation, it was essential that the degenerate codon at that position contain methionine to ensure the identification of the best variant. The efficient NDT codon was excluded for this reason. With a two site library, codon selection must also be efficient to reduce the screening effort. A two site NNK library would require the screening of 3066 colonies, which is near the maximum screening capacity of the high throughput assay. The DDK codon greatly reduced the screening requirement to 969 colonies, while still including methionine and maintaining a broad range of mutation (Table 4.3).

Table 4.3. Comparison of degenerate codon distributions of PheDH library 1.

Amino Acid	Codon DDK	Codon NDT	Codon NNK
Ala [A]			2
Arg [R]	1	1	3
Asn [N]	1	1	1
Asp [D]	1	1	1
Cys [C]	1	1	1
Gln [Q]			1
Glu [E]	1		1
Gly [G]	2	1	2
His [H]		1	1
Ile [I]	1	1	1
Leu [L]	1	1	3
Lys [K]	1		1
Met [M]	**1**		**1**
Phe [F]	1	1	1
Pro [P]			2
Ser [S]	1	1	3
Thr [T]			2
Trp [W]	1		1
Tyr [Y]	1	1	1
Val [V]	2	1	2
Stop	1		1
Codons	**18**	**12**	**32**
AA	**15**	**12**	**20**

4.4.1.5 Screening of LeuDH mutants and libraries

Each of the single mutants and combinations of these mutations were individually expressed and purified. Each variant was evaluated for activity toward 20 mM 1,3-DMBA using the 340 nm spectrophotometric assay (Section 3.2.3). The best combination of mutations was determined by the specific activity.

The best variant was further characterized for activity toward para-fluoro phenyl acetone (PFPA). PheDH's wild type keto-acid substrate is phenylpyruvate, and the

equivalent AmDH ketone substrate would be phenylacetone (Figure 4.5). Phenylacetone is not commercially available, because it is regulated as a Schedule 2 narcotic (113). This was overcome by the use of a fluorinated derivative, para-fluoro phenyl acetone (Sigma Aldrich) which is commercially available and unregulated.

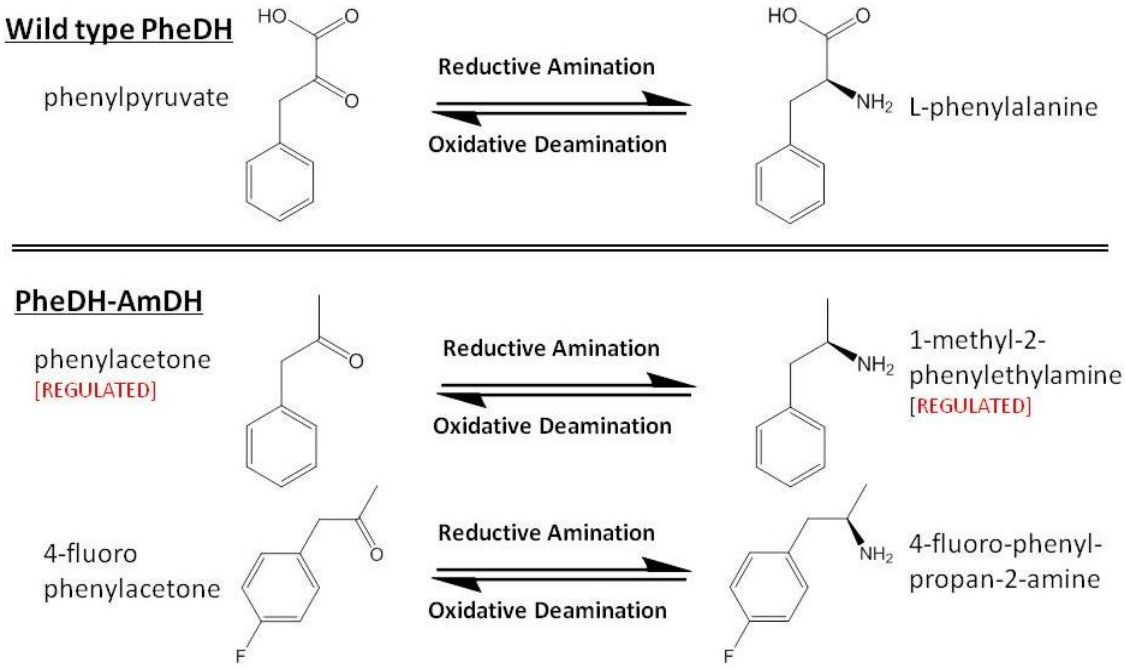


Figure 4.5. Wild type PheDH and PheDH-AmDH substrates.

The evaluation of the LeuDH-AmDH mutations on the PheDH scaffold identified the double variant K77M N276V as the most active, with a k_{cat} value of 0.128 s^{-1} and a reasonable K_M value of 4.61 mM (Section 4.4.2.2). Future variants with activity below this level were no longer considered significance. This elevated activity allowed for the application of a stringent screening assay (Figure 4.6) in screening the K77DDK N276DDK library. Screening in the deamination direction to exploit the higher activity was no longer

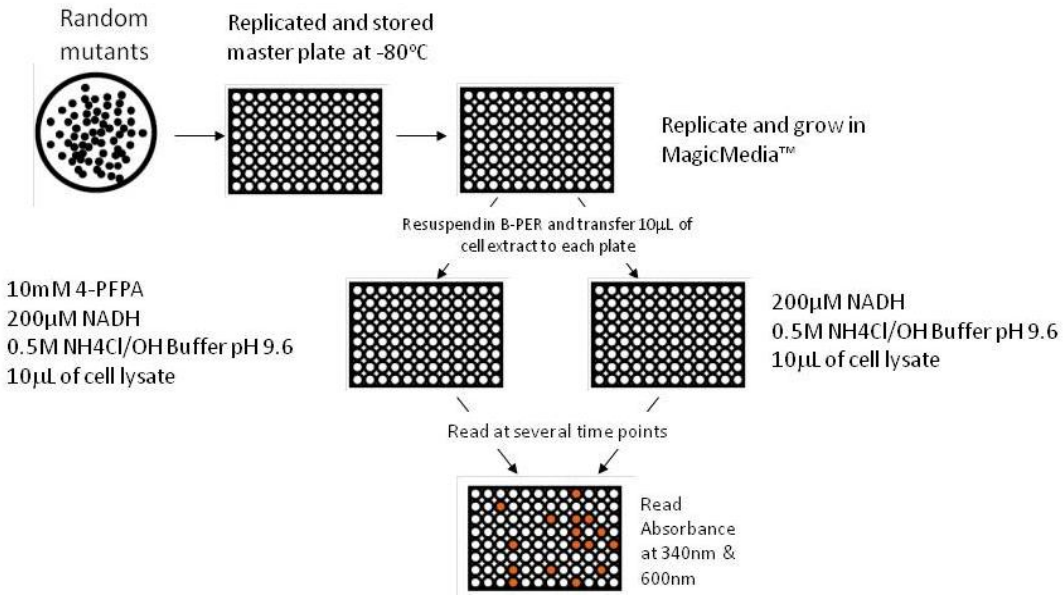


Figure 4.6. Forward screening assay applied to PheDH library 1 for reductive amination activity of *para*-fluoro phenyl acetone in 500 mM NH_4Cl/NH_4OH buffer pH 9.6, 200 µM NADH based upon 340 nm and 600 nm absorbance changes over multiple time points.

necessary. Instead, the amination activity was directly evaluated using a similar absorbance-based assay.

After mutagenesis, colonies were picked and expressed as described in Section 3.2.4. As the expression plates thawed, the cell pellets were gently vortexed with 30 µL of B-PER to uniformly resuspend and lyse the cells. The crude cell lysate was split in 10 µL aliquots into two plates; a reaction and background plate. A reaction mixture (200 µL) containing 10 mM PFPA and 200 µM NADH in 500 mM NH_4Cl/NH_4OH buffer was added to each well. The same mixture lacking PFPA was added to the background plate. Absorbance measurements at 340 nm and 600 nm began immediately unlike the previous end-point assays (Section 3.2.5), and continued periodically over the course of 1.5 hours.

Active variants were identified by the rate of change in absorbance at 340 nm, corresponding to the fastest activity. The wells exhibiting the fastest rate change over that of the background plate were selected for further characterization. Some variants consumed the cofactor too quickly (< 5 min) to be observed over multiple time points, and gave low absorbance in the initial time point. These variants were included for further characterization. The successful variants were sequenced (MWG Operon), and each unique sequence was expressed in pET28a/BL21 for histag purification and determination of k_{cat} and K_M values with PFPA.

4.4.2 Results and Discussion

4.4.2.1 Wild-type PheDH activity

The wild-type PheDH from *B. badius* was characterized to ensure proper expression and folding prior to mutation. The purified PheDH exhibited activity levels greater than 257 U mg^{-1} in the amination of phenylpyruvate (500 mM NH_4Cl/NH_4OH buffer pH 9.6, 200 μM NADH, 5 mM phenylpyruvate at 25 °C), and 39.5 U mg^{-1} in the deamination of L-Phe (0.1 M Glycine buffer pH 10.0, 1 mM NAD+, 20 mM L-Phe at 25 °C). K_M values were determined for the substrates when possible, and corresponded well to those values found in literature (110). Additionally, the K_I value of ammonia was determined to be 5.0 M. The K_M of phenylpyruvate could not be determined due to its high absorbance at the analytical wavelength. The wild-type enzyme was determined to be properly folded and active.

4.4.2.2 Comparison of beneficial mutations identified in LeuDH-AmDH

To directly and easily create a basis of PheDH-AmDH activity, each of the mutations indentified in LeuDH-AmDH were tested individually and in combination on the PheDH scaffold for the identification of residues which were influential in creating AmDH activity. The specific activities of the variants with 1,3-DMBA can be seen in Table 4.4.

The PheDH double variant, K77M N276V was the best combination of the LeuDH mutations. Both of these residues are known to interact with the wild-type substrate at the carboxyl moiety, and their synergistic effects are reflected in the large increase in AmDH activity when compared to their individual substitutions.

Table 4.4. AmDH activity on 1,3-DMBA and MIBK by PheDH variants derived from analogous LeuDH-AmDH mutations. Amination: 500 mM NH_4Cl/NH_4OH buffer pH 9.6, 200 μM NADH, 40 mM MIBK at 25 °C, Deamination: 0.1 M Glycine buffer pH 10.0, 1 mM NAD+, 20 mM 1,3-DMBA at 25 °C.

LeuDH Variant	**PheDH Variant**	**Deamination Activity 40 mM 1,3-DMBA , k_{cat} (s^{-1})**	**Amination Activity 40 mM MIBK, k_{cat} (s^{-1})**
Wild Type	Wild Type	---------- No Activity ----------	
Lys67Met	Lys77Met	0.0033 ±0.000	0.0009 ±0.000
Glu113Val	Thr123Val	0.0013 ±0.000	n.m.
Asn261Val	Asn276Val	0.0016 ±0.000	n.m.
Lys 67Met Glu113Val	Lys77Met Thr123Val	0.0010 ±0.000	n.m.
Lys67Met Asn261Val	**Lys77Met Asn276Val**	**0.0722** ±0.004	**0.0037** ±0.000
Lys67Met Glu113Val Asn261Val	Lys77Met Thr123Val Asn276Val	0.0011 ±0.000	n.m.

4.4.2.3 Screening of library and hits characterization

Prior to the screening of the PheDH K77DDK N276DDK library, the diversity of the mutations were checked. Twenty colonies were sequenced to ensure the desired breadth of mutation was achieved at each position, and a large percentage of parental variants were not present.

A total of 36 wells were labeled as highly active by the assay for subsequent characterization. Sequencing of these hits yielded 21 unique pairs of mutations, each of

Table 4.5. Randomly selected colonies of PheDH library 1 confirming diversity at positions K77 and N276 by successful application of overlap extension mutagenesis.

Colony	K77	N276
1	Leu	Trp
2	Asn	Phe
3	Val	Phe
4	Val	Asp
5	Arg	Gly
6	Arg	Asp
7	Asn	Lys
8	Gly	Ile
9	Tyr	Cys
10	Ser	Leu
11	Gly	Gly
12	Trp	Ile
13	Glu	Arg
14	Asp	Lys
15	Trp	Met
16	Glu	Lys
17	Glu	Glu
18	Cys	Cys
19	Gly	Leu
20	Lys	Trp

which were transformed into a pET28a plasmid for subsequent expression and his-tag purification. These purified proteins were individually characterized for activity toward the amination of PFPA to more accurately determine which combination of mutations performed best (Table 4.6). All 21 pairs of mutations were active amine dehydrogenases, eight of which provided k_{cat} values greater than 1 s^{-1} in the amination of PFPA. The top variant 1H4 contained the mutations K77S N276L with an apparent k_{cat} value of 2.8 s^{-1} and $K_{M,\ PFPA}$ value of 4.37 mM in 500 mM NH_4Cl/NH_4OH buffer pH 9.6 with 200 μM NADH. The K77W N276E variant also exhibited exceptionally high levels of activity with an apparent k_{cat} value of 2.63 s^{-1} and a $K_{M,\ PFPA}$ value of 5.32 mM. The K77S N276L variant was selected as the top candidate for further characterization, since it had the highest apparent k_{cat} value and a lower K_M than the K77W N276E variant.

Table 4.6. Characterization of histag purified PheDH library 1 hits. Kinetic parameters k_{cat} and K_M determined by non-linear fit for the amination of PFPA in 500 mM NH_4Cl/NH_4OH pH 9.6, 200 μM NADH.

Hit	K77 Residue	N276 Residue	Non-Linear Michaelis Menten k_{cat} (s^{-1})	K_M (mM)
1H4	**Ser**	**Leu**	**2.802 ±0.536**	**4.4 ±1.6**
7F10	**Trp**	**Glu**	**2.631 ±0.159**	**5.3 ±0.6**
4H6	Met	Met	1.818 ±0.190	7.9 ±1.3
7A3	Ser	Met	1.513 ±1.900	42.4 ±50.0
4D2	Ser	Ser	1.307 ±0.364	14.5 ±5.2
10F12	Cys	Leu	1.285 ±0.246	3.5 ±1.4
2F4	Met	Leu	1.283 ±0.100	4.2 ±0.6
3G7	Ser	Val	1.055 ±0.185	4.4 ±1.5
5A3	Met	Cys	0.896 ±0.080	1.5 ±0.4
8D10	Met	Leu	0.858 ±0.175	4.6 ±1.8
6D12	Ser	Phe	0.792 ±0.180	22.2 ±0.4
9G3	Met	Ser	0.661 ±0.074	4.9 ±1.0
10G6	Val	Leu	0.558 ±0.064	3.3 ±0.8
1E11	Cys	Ile	0.549 ±0.062	7.7 ±1.4
4A9	Gly	Ile	0.339 ±0.065	5.9 ±2.0
11E1	Ser	Gly	0.275 ±0.066	6.8 ±2.7
9A11	Gly	Cys	0.029 ±0.000	-
10A6	Ser	Glu	0.024 ±0.013	9.0 ±7.6
10H2	Trp	Ile	0.019 ±0.010	-
12D6	Trp	Gly	0.017 ±0.023	-
4C8	Gly	Val	0.017 ±0.008	-

4.5 Characterization of PheDH-AmDH

The K77S N276L variant was characterized in detail for a number of properties including; pertinent kinetic parameters, the breadth of substrate specificity, thermo-stability, overall conversion, and enantioselectivity.

4.5.1 AmDH activity and substrate specificity

Kinetic parameters, k_{cat} and K_M, were determined for the enzyme and each of the reductive amination substrates; respectively. The original assay conditions of 500 mM NH_4Cl/NH_4OH buffer did not saturate the enzyme with respect to NH_3. The k_{cat} value for this AmDH enzyme with all three substrates saturated in aqueous NH_4Cl/NH_4OH buffer at pH 9.6 was 6.85 s^{-1} at 25°C. This represents a near 15-fold enhancement above the maximum observed k_{cat} of 0.46 s^{-1} of the previously developed LeuDH-AmDH. The increase in buffer concentration did have a salt effect upon the K_M value of PFPA, which increased to 7.75 mM under saturated reaction conditions of 5 M NH_4Cl/NH_4OH buffer pH 9.6 with 200 μM NADH. The K_M value for NADH remained low as with the wild-type enzyme with a K_M value of 23.9 μM.

The significant increase in the K_M of NH_3 ($K_{M,\ NH3}$ = 1.27 M) seen in the conversion of para-fluoro phenyl acetone amination, as compared to the amination of phenylpyruvate ($K_{M,\ NH3}$ = 64.5 mM) can be attributed to decreased activation of the alpha carbon caused by changes in electron density with the new substrate. The original phenylpyruvate structure has a carboxyl group neighboring the alpha carbon where the nucleophilic attack of ammonia occurs. This carboxyl moiety will lessen the electron density at the alpha carbon make the nucleophilic attack more favorable compared to the ketone substrate.

The breath of the substrate profile was analyzed through the amination and deamination activity toward a number of different ketones and amines, respectively. The diversity of these ketones varied in structure from small aliphatic ketones such as 3-methyl-2-butanone, to larger aromatic ketones with additional functionality, such as phenoxy-2-propanone. Several other ketone compounds were found to exhibit reasonable levels of activity. (Table 4.7)

Table 4.7. Substrate specificity of top variant, K77S N276L PheDH. Amination in 500 mM NH_4Cl/NH_4OH, pH 9.6, 200 μM NADH, 20 mM substrate at 25 °C. [*] 40 mM, 20 mM of each enantiomer. [†] Deamination in 0.1 M glycine buffer, pH 10.0, 1 mM NAD+, 20 mM substrate at 25 °C.

Substrate	Activity (mU/mg)	Substrate	Activity (mU/mg)
R-MBA†	1.9 ±1.8	**cyclopentanone**	0.9 ±0.5
S-MBA†	< 1.0	**cyclohexanone**	27.5 ±1.4
R/S-MBA*,†	0.5 ±0.2	**ethylpyruvate**	18.4 ±5.9
MIBK	77.0 ±1.3	**benzaldehyde**	31.4 ±8.7
1,3-DMBA†	166.3 ±0.0	**2-methylcyclohexanone**	19.3 ±1.1
acetophenone	< 1.0	**3-methylcyclohexanone**	41.1 ±2.1
phenoxy-2-propanone	540.8 ±6.9	**3-methyl-2-butanone**	72.7 ±1.4
2-hexanone	155.7 ±1.4	**1-Boc-piperidone**	7.8 ±1.1
3-hexanone	1.6 ±0.5	**Benzyloxyacetic acid**	22.6 ±0.7
3-pentanone	1.3 ±0.0		

The enzyme shows increased activity toward methyl-ketones versus those which are ethyl-ketones or cyclic-ketones. The top five most active ketones; PFPA, phenoxy-2-propanone, 2-hexanone, methyl isobutyl ketone and 3-methyl-2-butanone are all methyl-

ketones. This specificity is also observed in the large differences in activity toward 2-hexanone (155.7 mU/mg) versus 3-hexanone (1.6 mU/mg). Despite the structural similarities, the methyl-ketone 2-hexanone, exhibits approximately 100-fold higher activity. Similarly, the enzyme requires at least a one carbon linkage between the alpha carbon and the substrate's phenyl ring. This is reflected by the unobservable activity toward acetophenone, as well as the low deamination activity of methyl benzyl amine.

4.5.2 Thermostability

The thermostability of the AmDH was compared to its wild-type scaffold using two different methods, circular dichroism (CD) and a temperature versus activity profile. The mutation of the enzyme scaffold has affected the thermostability, in addition to the enzyme's function. Both the wild-type PheDH and the AmDH exhibit similar characteristics in the loss of secondary structure as indicated through CD. Both enzymes showed the same gradual loss of structure at temperatures up to 30 degrees Celsius (Figure 4.7 and 4.8), implying that the cause for the early loss of structure is inherent in the wild-type protein. Beyond 30 degrees, there is an earlier onset in the pronounced loss of structure of the double variant compared to the wild type; however, the slope folded fraction over temperature stays constant, indicative of constant enthalpy of melting. The K77S N276L mutations have slightly destabilized the protein structure, resulting in a melting point of 59.9°C, 4.4°C less than the wild type. The specific cause for this change is difficult to determine. It likely to be either disruption of the 'ideal' packing of the wild-type enzyme's hydrophobic core or a net decrease in hydrophobicity by the amino acid substitutions which leads to a lower folding entropy (114, 115).

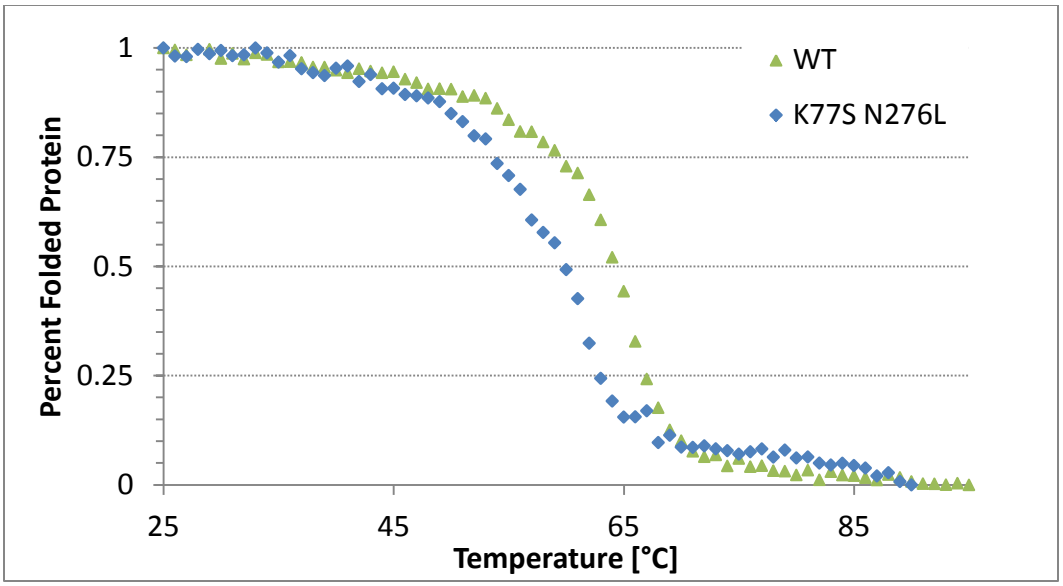


Figure 4.7. Fraction of folded PheDH/AmDH as measured by circular dichroism spectroscopy of wild type PheDH and top PheDH-AmDH variant, K77S N276L with normalized ellipticity representing the percentage of folded protein (protein concentrations 100 μg mL^{-1} in 50 mM sodium phosphate buffer pH 8.0).

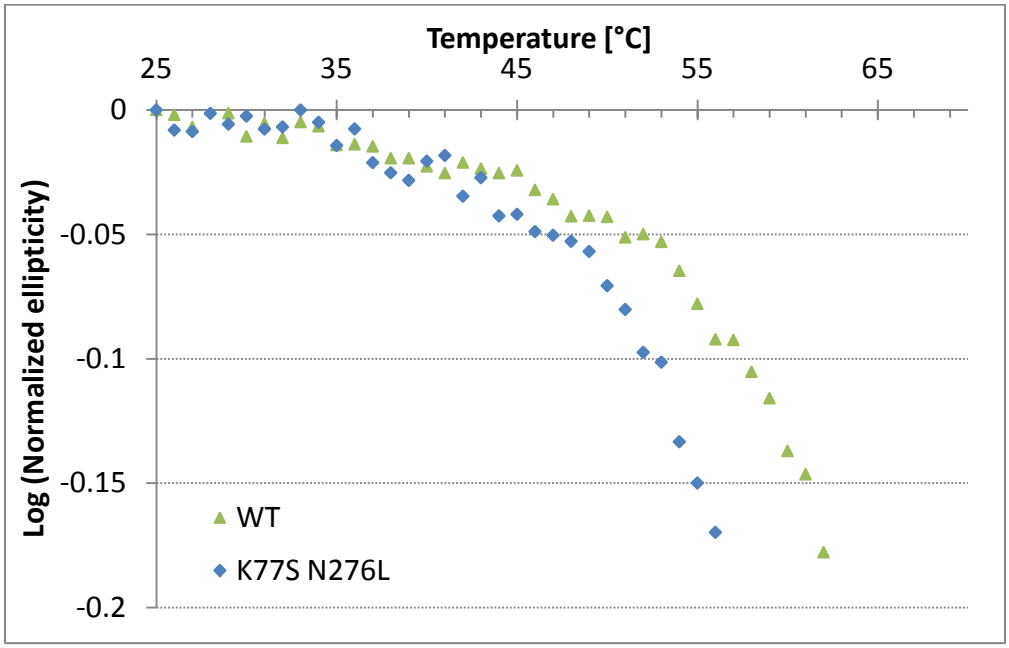


Figure 4.8. Semi-log plot of fraction of folded PheDH/AmDH as measured by circular dichroism spectroscopy of wild type PheDH and top PheDH-AmDH variant, K77S N276L with normalized ellipticity representing the percentage of folded protein (protein concentrations 100 μg mL^{-1} in 50 mM sodium phosphate buffer pH 8.0).

The thermostability of the double variant K77S N276L was additionally determined by a temperature versus activity plot (Figure 4.9). This plot shows the enzyme's Arrhenius activation through the lower temperature range (10-45 °C), although this thermal activation should appear to be exponentially increasing (116). The near linear increase corresponds to a low activation energy of 241 J mol^{-1}. The maximum specific activity was 11.6 U mg^{-1} at a temperature of 50 °C. Above 50 °C, the enzyme began to rapidly lose activity as a result of denaturation. This deactivation correlated well with the circular dichroism data, where a dramatic loss of secondary structure was observed from 50 °C to 70 °C.

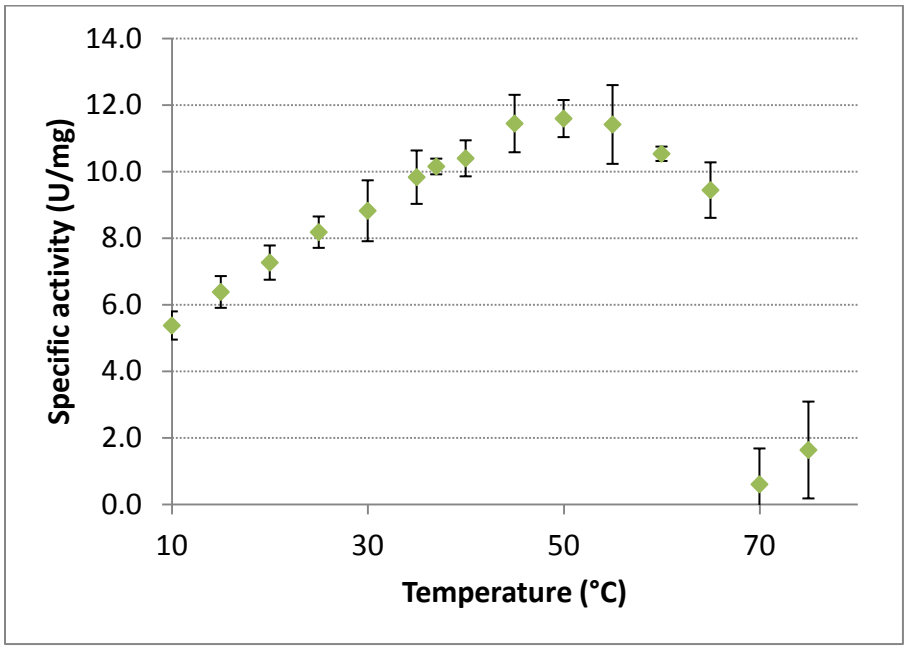


Figure 4.9. Temperature vs. Activity of PheDH K77S N276L double variant. Amination activity measured in 5 M NH_4Cl/NH_4OH buffer pH 9.6, 200 μM NADH, 20 mM PFPA for 2 minutes.

4.5.2 Conversion with cofactor recycle system

A reaction system capable of high-level conversions was demonstrated by pairing the AmDH with a GDH cofactor recycle system. The GDH consumes glucose present in solution to cheaply and efficiently regenerate the NADH cofactor, allowing conversion to continue beyond stoichiometric depletion of NADH. In a 100 mL reaction volume containing 20 mM PFPA, conversion was allowed to continue for 48 hours followed by extraction and ^{1}H-NMR to confirm the product formation and conversion levels. The addition of 7.48 mg of the AmDH enzyme converted 89.4% of the 4-fluorophenyl acetone substrate to the product 1-(4-fluorophenyl)-propane-2-amine over 48 hours with an isolated yield of 66.2%. Under the same reaction conditions, 17.46 mg of the AmDH enhanced the overall conversion to 93.8% and 73.9% yield. Evidence of chirality was indicated by different NMR signals of the C3 hydrogens (δ 2.67, 2.49) which are adjacent to the C2 stereocenter (117) (Figure 4.10).

4-fluorophenyl acetone substrate;

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.26-7.00 (m, 4H), 3.68 (s, 2H), 2.16 (s, 3H).

(*R*)-(-)-1-(4-fluorophenyl)-propane-2-amine product;

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.26-6.95 (m, 4H), 3.12 (sex, 1H, J = 3.9 Hz), 2.67 (dd, 1H, 6.3 Hz, 13.2 Hz), 2.49 (dd, 1H, 6.3 Hz, 13.2 Hz), 1.10 (d, 3H, 6.4 Hz).

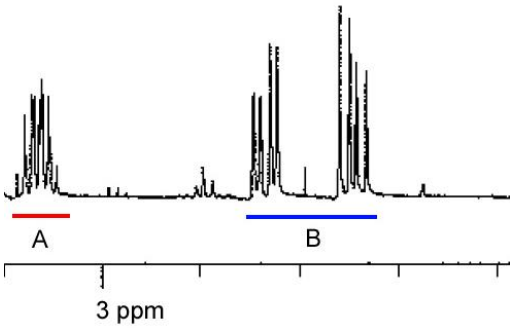


Figure 4.10. Evidence of chiral 1-(4-fluorophenyl)-propane-2-amine via ^{1}H NMR. A) Sextuplet splitting representative of the proton at the chiral center. B) A pair of doublet-doublet splits from the C3 hydrogens reflecting the chiral center.

4.5.3 Enantioselectivity

Equally important to conversion is maintaining the selectivity of the enzyme after altering the substrate specificity. Preservation of the wild type (*S*)-selectivity towards phenylalanine will result in the asymmetric production of (*R*)-1-(4-fluorophenyl)-propyl-2-amine due to a change in Cahn-Ingold-Prelaw priority. The product amine resulting from enzymatic conversion with the co-factor recycle was isolated and derivatized using trifluoroacetic anhydride for analysis via chiral gas chromatography. An enantiomeric excess of > 99.8% toward (*R*)-1-(4-fluorophenyl)-propyl-2-amine was confirmed, however, further resolution of the selectivity was limited by the separation capability of the chiral column (118). Throughout all experiments, the (*S*)-enantiomer remained below the detection limits. Even as the enzyme approaches levels of conversion beyond 90%, the high enantioselectivity toward (*R*)-1-(4-fluorophenyl)-propyl-2-amine is maintained. This renders the AmDH/GDH system a viable method for the production of chiral amines.

Confirmation of the (*R*)-selectivity was achieved through polarimetry, since single enantiomers of (4-fluorophenyl)-propyl-2-amine are not commerically available. Suspension of the isolated product amine in methanol resulted in a negative optical rotation (-1.73°, 10 cm path length at 0.3 mM amine), confirming the (*R*)-(-)-enantiomer to known standards (119).

4.6 Conclusions

Previous knowledge in the evolution of an amine dehydrogenase, particularly the influence of binding pocket residues Lys77 and Asn276, allowed for a direct route to an active AmDH from PheDH. With a reasonably active starting variant, a stringent high-throughput screening of a single two-site library further enhanced the desired reductive amination activity to a k_{cat} value of 6.85 s^{-1} at 25 °C. The high enantioselectivity was maintained throughout the evolution process. The enzyme's *e.e.* value of > 99.8% toward the (*R*)-enantiomer, even after high levels of conversion, renders it an attractive candidate for the asymmetric production of chiral amines. The reaction can be simply and cheaply driven to conversions in excess of 90% when paired with a cofactor recycle system. Despite a 4.4 degree Celsius decrease in thermostability, the enzyme remains broadly activity over a large temperature range from 10 to 50 °C, with a maximum observed specific activity of 11.6 U mg^{-1}. The novel AmDH exhibits amination activity towards a range of ketone substrates, which makes it a good starting point for further evolution to increase activity toward specific API targets.

CHAPTER 5: BROADLY APPLICABLE MUTATIONS TO CREATE AMINE DEHYDROGENASE ACTIVITY

5.1 Introduction

The previous development of two AmDHs (Chapters 3 and 4) has allowed for the novel conversion of a prochiral ketone and free ammonia to a chiral amine product. These enzymes open the possibility of new synthetic pathways and could be evolved toward target substrates. The availability of two scaffolds significantly expands the substrate specificity, especially with the large differences in binding pocket size and substrate functionality between LeuDH-AmDH and PheDH-AmDH. During the evolutionary route of producing these enzymes, the assays concurrently identified the mutations K67S and N261L (LeuDH nomenclature, and equivalent PheDH residues) as essential for amination activity. The mutation of these residues is believed to be broadly applicable across the Glu-Leu-Phe-Val dehydrogenase sub-family of enzymes. To test this hypothesis, a third scaffold from ValDH will be evolved at its analogous residues for AmDH activity, potentially giving credence to a family of AmDHs.

5.2 Materials and Methods

5.2.1 Selection of valine dehydrogenase from Glu-Leu-Phe-Val family

The Glu-Leu-Phe-Val dehydrogenase sub-family (Accession number: cd05211) contains a large number of characterized proteins which are further divided into smaller sub-families. The *B. stearothermophilus* LeuDH, *Rhodococcus sp.* M4 PheDH and *B. badius* PheDH all exist within this large sub-family, and are even further grouped in a smaller sub-family NAD_bind_Leu_Phe_Val_DH (Accession number: cd01075). This

smaller sub-family grouping is a result of their highly conserved binding motifs of the NADH cofactor. Glutamate dehydrogenases (GluDH) are not a part of this smaller family, but are grouped into two separate GluDH subfamilies (Accession number: cd01076 and cd05313) within the greater Glu-Leu-Phe-Val dehydrogenase sub-family. Consequently, ValDH exhibits a much higher sequence identity to PheDH and LeuDH (typically greater than 60%), and GluDHs are more distant relatives (typically near 30% identity). The higher identity will increase the likelihood of successful translation of AmDH mutations to the ValDH scaffold. ValDH was chosen as the scaffold for demonstration of the broad application of AmDH mutations.

Three ValDHs were selected as potential scaffolds due to their previous characterization in literature. All three proteins hail from the *Streptomyces* genus of Actinobacteria, *Streptomyces coelicolor* (120-122), *Streptomyces fradiae* (123-125), and *Streptomyces cinnamonensis* (122, 126, 127). All three species suffer from high guanine and cytosine (GC) content which can cause difficulty in PCR. ValDH from *Streptomyces cinnamonensis* was ultimately chosen as the scaffold, since it had been previously been overexpressed in a pET vector, *E. coli* BL21(DE3) expression system resulting in large amounts of soluble and functional protein in the cytoplasm (126). The resulting ValDH-based AmDH (ValDH-AmDH) would convert 2-methyl-3-butanone to chiral 1,2-dimethylpropylamine (Figure 5.1)

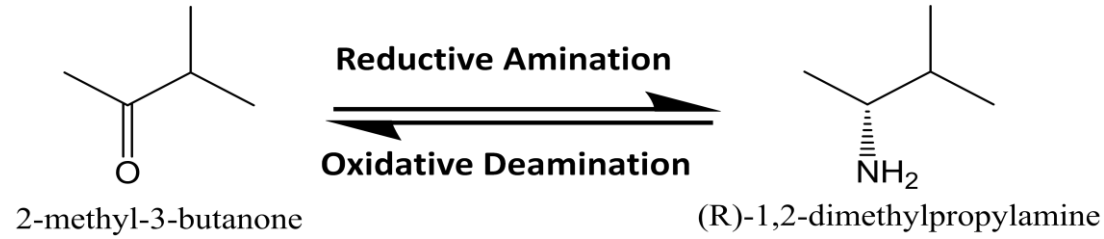


Figure 5.1. Desired reductive amination activity of ValDH-AmDH converting 2-methyl-3-butanone to (*R*)-1,2-dimethylpropylamine.

5.2.2 Codon optimization of valine dehydrogenase *Streptomyces cinnamonensis*

The usage of unfavorable codons in a highly expressed gene in *E. coli* can decrease the translational rate (128, 129). With redundancy present in the genetic code for most amino acids, the unfavorable triplets can be replaced with more favorable codons representing the amino acid. For example, arginine codons AGA and AGG have a usage frequency below 0.2% are replaced with commonly used codons CGT and CGC. With these codons removed from the gene sequence prior to synthesis, the resulting construct will exhibit improved expression over the wild-type sequence while still encoding the identical peptide. Additionally, long stretches of consecutive GC-base pairs can cause difficulties in polymerase chain reaction (PCR) amplification of DNA. The ValDH gene from *S. cinnamonensis* has a GC-content of 71% which includes several long stretches of consecutive GC-base pairs as long as 19-base pairs. Codon selection was also optimized to remove these GC-rich regions from the gene and reduce the overall GC-content to 59%. The optimized gene sequence can be found in Appendix A5.2.

5.2.3 Gene synthesis, preparation, overexpression, and purification

The codon-optimized ValDH gene was synthesized by MWG Operon. The synthesis included the insertion of *Nde*I and *Hind*III restriction sites prior to and following the designed gene for simple digestion and ligation into pET28a plasmid.

The plasmid was prepared, transformed and expressed using the protocol described in Section 3.2.1. Expression at 30 °C, instead of 37 °C in auto-inducing MagicMedia™ was the only deviation from the described protocol. Protein was purified using the histag purification protocol described in Section 3.2.2.

5.2.4 ValDH-AmDH K76DDK N273DDK library

The ValDH-AmDH library will be generated using an analogous protocol to that described in Section 3.2.4 by applying the primers described in the Table 5.1. The resulting screening requirement of the two site (K76DDK, N273DDK) library is 969 colonies.

Table 5.1. Mutational primers applied to *S. cinnamonensis* ValDH to create two-site DDK ValDH-AmDH library.

ValDH Residue		Primer	
K76	Fwd	5' CTGAGCCGTGGCATGAGCTAT**DDK**AACGCGATGGCGGGTCT	3'
	Rev	5' AGACCCGCCATCGCGTT**MHH**ATAGCTCATGCCACGGCTCAG	3'
N273	Fwd	5' TGTGCGGTGCGGCAAAC**DDK**CAGCTGGCGCATCCGG	3'
	Rev	5' CCGGATGCGCCAGCTG**MHH**GTTTGCCGCACCGCACA	3'

The resulting library will be picked and expressed according to Section 3.2.4, at a temperature at 30 °C. After checking the libraries diversity, it will be screened in the

same manner as the previously described PheDH-AmDH library (Chapter 4, Figure 4.6) with the substrate as the only difference. Instead of PFPA, 20 mM 2-methyl-3-butanone (2M3B) will be the ketone analog closest to the wild type keto-acid.

5.3 Results and Discussion

5.3.1 Characterization of wild-type ValDH activity

The wild-type ValDH was expressed to ensure proper expression and folding prior to mutation. Histag purification yielded functional ValDH protein at a concentration of 229 μg mL^{-1}. The successful purification can be seen in the strong enrichment and isolation of a 43.3 kDa protein via SDS-PAGE (Figure 5.2). The resulting protein was screened for deamination of L-valine in a 250 mM sodium carbonate buffer, pH 10 containing 1 mM NAD+. The specific activity, 16 U mg^{-1} matched the literature values for deamination (127), confirming the presence of functional protein.

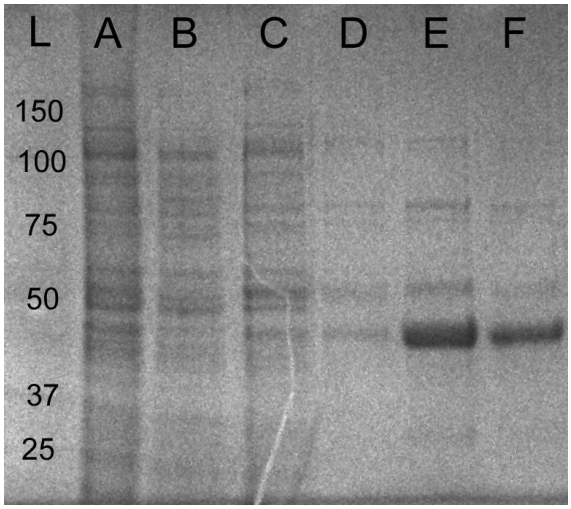


Figure 5.2. Histag purification of wild type *S. cinnamonensis* ValDH. Lane L: Ladder, A: Clarified cell lysate, B: Unbound clarified cell lysate, C: 50 mM imidazole wash 1, D: 50 mM imidazole wash 2, E: Pure protein elution-overloaded, F: Pure protein elution-standard loading.

5.4 Conclusions

Valine dehydrogenase from *Streptococcus cinnamonensis* was chosen from a subset of well characterized ValDHs as an example scaffold for the broad application of mutations K67S and N261L (LeuDH-AmDH). The ValDH gene has been codon-optimized to avoid prolonged stretches of consecutive GC-base pairs, high overall GC-content, and usage of rare codons. The optimized gene has been overexpressed in *E. coli* BL21 (DE3) competent cells and successfully purified over a Ni^{2+}-column via IMAC using an N-terminal hexa-histag. The resulting protein exhibited deamination activity levels similar to those reported in literature (16 U mg^{-1}). The gene is prepared for mutagenesis generating library 1; a broad two-site library over the most influential residues for AmDH activity. Hits indicated by the absorbance based assay will be further characterized for detailed kinetic parameters and substrate specificity.

CHAPTER 6: RECOMMENDATIONS AND CONCLUSIONS

6.1 Recommendations

6.1.1 Further protein engineering of PheDH- and ValDH-AmDHs

Directed evolution has been successfully applied in many instances to improve the catalytic properties of enzymes (30, 58). The PheDH and ValDH scaffolds have not yet been subjected to the evolutionary process. Only the single PheDH library of residues K77DDK and N276DDK has been screened. There is still a large potential for identification of beneficial mutations by screening other residues of the scaffold. Slight adjustments in the distances between catalytic residues can have a strong effect on reaction rates (130, 131). The first-shell residues of the active AmDHs are responsible for both direct interactions with active-site residues and binding of the substrate, and should be evolved using a process similar to that applied to LeuDH-AmDH (Table 6.1). Optimization of these protein-substrate interacts should result in increased catalytic constant k_{cat} and a lower K_M value.

The residues A38 and F137 (*Rh. sp.* M4 PheDH) are of particular interest due to their interactions with the *para*-substituted region of the PheDH substrate (Figure 6.1). The *para*-fluoro substitution of phenyl acetone creates not only a larger, but more polar moiety at this position. Substituting the A38 with a more polar side chain, or decreasing the size of the F137 may lead to more favorable substrate binding.

Evaluation of these selected amino acids will allow us to further increase the amine dehydrogenase activity, while significantly reducing the sequence space by focusing the mutagenesis to target, binding-pocket residues.

Table 6.1. Alignment of binding pocket residues for *B. stearothermophilus* LeuDH, *Rhodococcus sp.* M4 PheDH, *B. badius* PheDH, and *S. cinnamonensis* ValDH.

LeuDH	**PheDH**		**ValDH**	
B. stearothermophilus	Rhodococcus M4	Bacillus badius	Strep. Cinnamonensis	**Rationale for mutation**
L39	A38	L49	L48	Residues interact with the *para*-substituted region of the Phe substrate
G40	G39	G50	G49	
G41	G40	G51	G50	
M64	M63	M74	M73	Side chain interactions with neighboring Lysine
K67	**K66**	**K77**	**K76**	Direct interaction with the carboxyl of the keto-acid substrate, along with Asn.
A112	G116	G122	A121	Residues interact with the posterior portion of the substrate, opposite the reacting ketone/amine.
E113	**P117**	**T123**	**C122**	
D114	D118	D124	D123	
V115	V119	M125	V124	
T133	F137	V143	T142	Residues interact with the *para*-substituted region of the Phe substrate
N261	**N263**	**N276**	**N273**	Direct interaction with the carboxyl of the keto-acid substrate, along with Lys.
V290	**A292**	**L305**	**V302**	Residues shape substrate binding pocket
V293	L295	V308	V305	

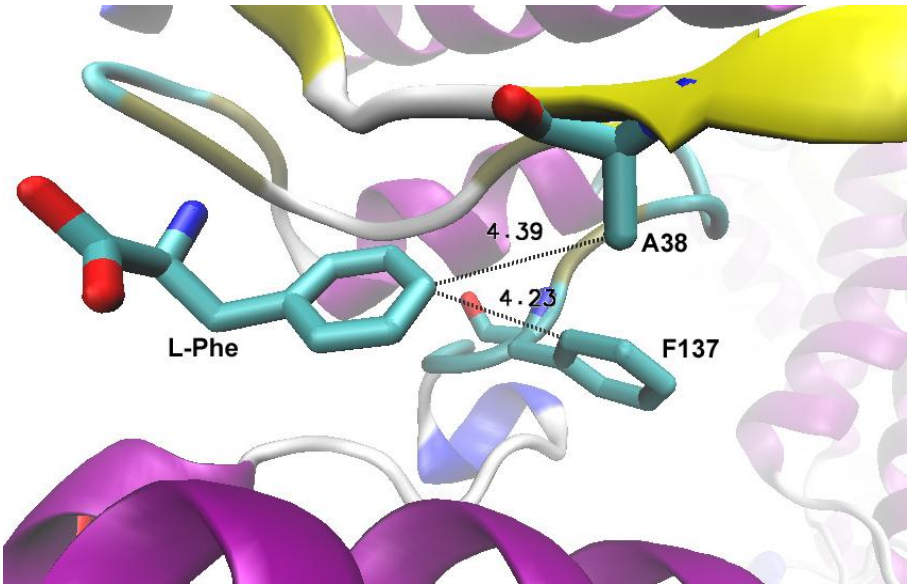


Figure 6.1. Binding pocket of *Rhodococcus sp.* PheDH with bond distances in Å between the L-Phe substrate C4 carbon and neighboring side chains of A38 and F137.

6.1.2 Stabilization of AmDH scaffolds

To achieve a robust industrial enzyme process, high thermostability and activity of the enzyme catalyst are required. Numerous enzymes have been evolved to improve thermostability and tolerance to nonnative process conditions through a variety of strategies. These strategies include introduction of disulfide bonds (132), improved core packing (133-136), increased rigidity (54, 115), and optimized surface charges (137, 138).

Several online tools allow for the prediction of potential disulfide bond pairs, i.e. SS-BOND or GDAP (139, 140). An example analysis of the *Rhodococcus sp.* M4 PheDH using SS-BOND can be found in Appendix A.3. This algorithm approximates the location of β-carbon atoms from the position of the backbone, and screens this set for suitable β-carbon to β-carbon distances and bond angles favorable for a disulfide bond. The final confirmations are energy minimized and reported in the output. These results

can then be further screened experimentally or computationally for their effect on protein stabilization.

The tight packing of the protein's core is essential for stability, and the elimination of voids in the protein structure results in a decreased free energy. Recently developed software, RosettaHoles2 and Rosetta$_{VIP}$ allow for the quantitative and visual assessment of the packing in the protein core (135, 136, 141). These flaws in core packing can then be improved by directed evolution to create more energetically-favored folded protein states.

6.1.3 Formate dehydrogenase cofactor recycle systems

The application of NAD(P)H-dependent dehydrogenases, such as the aforementioned AmDHs, require the use of a cofactor regeneration system. The regeneration of the cofactor significantly reduces the substrate cost by requiring only catalytic amounts of the expensive cofactor, instead of large stoichiometric quantities. Glucose dehydrogenase (GDH) has been successfully demonstrated to regenerate the cofactor for the LeuDH-AmDH (79), but a more ideal cofactor recycle system would be formate dehydrogenase (FDH). FDH would allow for the same regeneration of cofactor, and simultaneously allow for improved atom economy (Table 6.2). Instead of supplying the additional glucose substrate for the GDH system, the formate dehydrogenase could rely on formate present in an ammonium formate buffer system. Formate dehydrogenases offer the additional advantage of having specific activities more closely to that of the AmDHs, approximately 10 U mg^{-1} simplifying the use of co-expression (142).

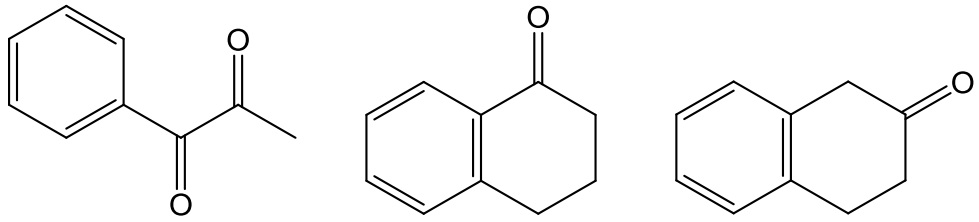
Table 6.2. Mass efficiency estimation for GDH and FDH cofactor recycle systems. A 1 L-basis was chosen with 100 mM substrate concentration and 92% conversion to product.

GDH-based cofactor recycle			
Chemical	**Molarity (M)**	**MW (g mol^{-1})**	**Mass (g)**
NH_4Cl buffer	0.5	59.5	29.75
NADH	0.0002	763	0.15
MIBK	0.1	100.2	10.02
Glucose	0.1	180.2	18.02
Total			57.94
Product produced (92% conversion)			
1,3-DMBA	0.092	101.2	9.31
Estimated mass efficiency			**16.1%**

FDH-based cofactor recycle			
Chemical	**Molarity (M)**	**MW (g mol^{-1})**	**Mass (g)**
NH_4HCO_2 buffer	0.5	63.1	31.55
NADH	0.0002	763	0.15
MIBK	0.1	100.2	10.02
Total			41.72
Product produced (92% conversion)			
1,3-DMBA	0.092	101.2	9.31
Estimated mass efficiency			**22.3%**

6.1.4 Organic co-solvent systems

Enzymatic synthesis in non-aqueous media is an attractive alternative to traditional chemical synthesis (143). Many target active pharmaceutical intermediates, such as tetralones or diketones (Figure 6.2), have limited solubility in aqueous media. To increase the solubility, these substrates require the application of an organic cosolvent. Solvent selection is important and can greatly impact the hazard, environmental-friendliness, and cost of synthesis (144). Ideally, the organic solvent would comply with ACS GCI Pharmaceutical Roundtable Solvent Selection Guide to maintain the green benefits of an enzymatic reaction (145). Amino acid dehydrogenases have only been sparsely studied in organic solvents (146), but their successful application will allow for the more hydrophobic substrates.



1-phenyl-1,2-propanedione alpha-tetralone beta-tetralone

Figure 6.2. Examples of hydrophobic substrates for ketone to amine reduction in organic media.

Preliminary experiments with PheDH-AmDH K77M N276V have shown activity with an acetone cosolvent. The enzyme tolerates acetone concentrations as high as 20% v/v. The PFPA substrate has a limited solubility of 40 mM in aqueous media, but was increased to 60 mM with the addition of acetone. The increased substrate concentration resulted in a 40% increase in the maximum observable k_{cat} (Figure 6.3), from 0.11 s^{-1} in aqueous solution to a value of 0.26 s^{-1} with 20% acetone. The K_M were affected by the addition of acetone, increasing from a value of 4.6 mM to 23.1 mM in aqueous and 20% acetone, respectively. Further evaluation of organic solvents, combined with evolution to solvent stability (146) will allow for further increases in conversion rates and specificity toward hydrophobic substrates.

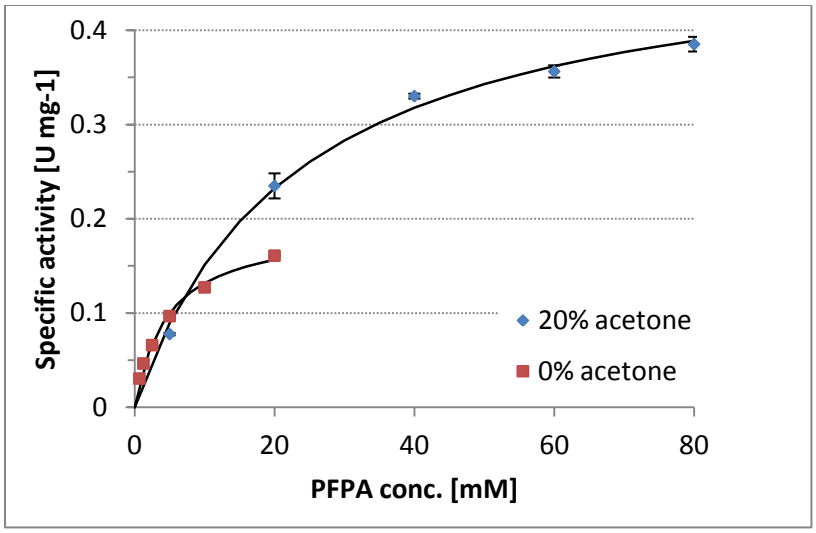


Figure 6.3. PheDH-AmDH K77M N276V observed rate enhancement with 20% acetone by elevating substrate concentration.

6.1.5 Evolution toward pharmaceutical targets

The ability to alter a protein's sequence, structure and function has allowed scientists to evolve proteins toward new functions. Novel function can be achieved by directed evolution, but this approach must be focused toward a particular substrate to achieve the desired specificity. As stated by Tawfik *et al.*, the first rule of directed evolution is 'you get what you select for' (49). Directed evolution of the same residues screened to create the original AmDH, have the potential to create analogous amination of other ketones as achieved with ω-transaminases (17).

6.1.6 Scale up and whole-cell systems for target substrates

Industrial application of biocatalysts requires the scale-up of enzymatic reactions from lab-scale to pilot and industrial scale. Scientific and economic considerations must be accounted for during catalyst application on an industrial scale. A lab-scale biocatalytic system which probes economic feasibility of an industrial scale process for

the novel amine dehydrogenase system can be achieved through two routes; i) whole-cell catalysis containing a suitable co-expression system capable of cofactor recycle, ii) an enzyme membrane reactor capable of continuous production of chiral amine and co-factor recycle.

With nearly equivalent specific activities, the AmDH (1-10 U mg^{-1}) and formate dehydrogenase (FDH) from *Candida boidinii* (6 U mg^{-1}) can simply be expressed on a dual gene plasmid, such as pETduet-1 (Merck Millipore, Billerica, MA) in *E. coli* BL21 (DE3) competent cells. The resulting co-expression should give similar specific activities per gram of soluble protein allowing for successful catalysis. However, if future AmDHs have substantially higher activities a two plasmid approach may be required. This approach has been successfully applied to the whole-cell asymmetric reductive amination of α-keto acids which represents a useful methodology for the synthesis of α-amino acid L-leucine (147). LeuDH from *B. cereus* was chosen for the reductive amination and FDH from *C. boidinii* for the cofactor regeneration. The main challenge for successful co-expression of this system was the large discrepancy of specific activity between the two enzymes; ~400 U/mg for LeuDH and only 6 U/mg for FDH. This obstacle was overcome creating a co-expression system based on the same inducible promoter, but locating the genes on two *E. coli* plasmids with different copy numbers. To compensate for its low activity, FDH was expressed in plasmid pAM3-25 which has a higher copy number. Conversely, LeuDH was expressed in plasmid pAM10-1, a medium copy number plasmid (148). The two plasmids were transformed into BW3110, an *E. coli* strain suitable for high-density fermentation (149). The resulting combination gave similar

specific activities per gram of soluble protein and in specific cases achieved nearly 100% conversion to L-leucine.

An alternative option to whole-cell catalysis is an enzyme membrane reactor (EMR); which have been successful in the industrial application of enzymatic processes. In these systems, enzymes do not require simultaneous expression and instead can be combined at the appropriate ratio within the reactor. EMR reactors do however require purification of the desired enzymes, which can be achieved through various purification techniques (8). This system would require use of the same AmDH and FDH enzymes as the whole-cell system. Each enzyme would be independently expressed and loaded to the reactor to yield similar specific activities per gram of protein. Reaction conditions such as flow rate, temperature, substrate concentration, and enzyme loading are optimized to increase conversion. Enzymes can easily be contained within the reactor by restriction of the membrane pore size. The retention of the cofactor poses a more difficult problem. To keep the cofactor from penetrating the membrane, it may be enlarged with polyethylene glycol (PEG) (150). Otherwise, it can simply be recharged along with substrate. Lastly, this finalized reaction process would be used to isolate the reaction product (~ 10g) for analysis of chemical and enantiomeric purity.

6.2 Conclusions

This work describes the development of a novel class of amine dehydrogenases for the synthesis of chiral amines. Directed evolution has successfully enhanced the binding and activity for an unnatural ketone substrate. Eleven rounds of directed evolution have completely altered the LeuDH enzyme's specificity and created amination activity. Each round of mutagenesis focused on a region of the binding pocket, including

simultaneous mutation of neighboring residues to capture synergistic effects. These variants were screened by various high-throughput assays identifying minute increases in amination activity, and successful mutations were carried into future rounds of mutagenesis.

The largest improvements in activity were achieved by the cooperative mutation of residues K67 and N261. Their simultaneous mutation ultimately identified the most active quadruple mutant, K67S/E113V/N261L/V290C, with novel reductive amination activity of 0.69 U mg^{-1} with a corresponding k_{cat} value of 0.46 s^{-1}. Within the first round of mutation, the native activity toward L-Leu of 112 U mg^{-1} (98) was eradicated to less than 2 mU mg^{-1}; completely altering the enzyme's specificity. The enantioselectivity of the wild-type enzyme was maintained despite the drastic changes to the binding pocket and yielded (*R*)-1,3-DMBA with an *e.e.* value of 99.8% at 92.5% conversion. This LeuDH-AmDH exhibited activity toward a number of different substrates. The enzyme has also maintained its wild type stability making it an attractive catalyst in the synthesis of chiral amines. This was the first example of a cofactor-dependent amine dehydrogenase capable of selectively synthesizing chiral amines from a prochiral ketone and free ammonia.

The two most influential mutations in the LeuDH-AmDH were applied to second scaffold, PheDH from *B. badius*. A single two-site library, K77DDK N276DDK, directly identified a double variant with significantly increased AmDH activity. This double variant K77S N276L exhibited a k_{cat} value of 6.85 s^{-1} in the reductive amination of PFPA to (*R*)-(+)-1-(4-fluorophenyl)-propane-2-amine. The high selectivity of the wild type enzyme was preserved throughout the evolution process. The PheDH-AmDH resulted in

an *e.e.* of > 99.8% toward the (*R*)-enantiomer, even after high levels of conversion, rendering it an attractive candidate for the asymmetric production of chiral amines. The reaction can be simply and cheaply driven to conversions in excess of 90% when paired with a cofactor recycle system. Despite a 4.4 degree Celsius decrease in thermostability, the enzyme remains broadly activity over a large temperature range from 10 to 50 °C, with a maximum observed specific activity of 11.6 U mg^{-1} (50 °C, and a pH value of 9.6). The novel AmDH exhibits amination activity towards a range of ketone substrates, which makes it a good starting point for further evolution to increase activity toward specific API targets.

The applicable breadth across the Glu-Leu-Phe-Val dehydrogenase sub-family of the two key mutations (PheDH: K77 N276) observed in creating AmDH will be tested upon a third scaffold, ValDH from *Streptomyces cinnamonensis*. ValDH exhibits many of the same folding and mechanistic features as the previously developed AmDHs. Through a similar application of directed evolution to the PheDH scaffold, a focused library of mutants will be screened for novel amination activity on the ValDH. The corresponding gene has been codon-optimized and prepared for expression in *E. coli*.

Future works in the development of amine dehydrogenases should be targeted to the industrial application of these enzymes. Activities can possibly be enhanced further through directed evolution to make for even more attractive industrial catalyst. Several other aspects pertaining to industrial application of a catalyst have yet to be addressed and should be the priority of future experiments. Stabilization, co-solvent systems and process scale up have the potential of resulting in a robust catalyst applicable to the asymmetric production of target APIs.

APPENDIX

Appendix A.1 Glu-Leu-Phe-Val sub-family phylogenic tree

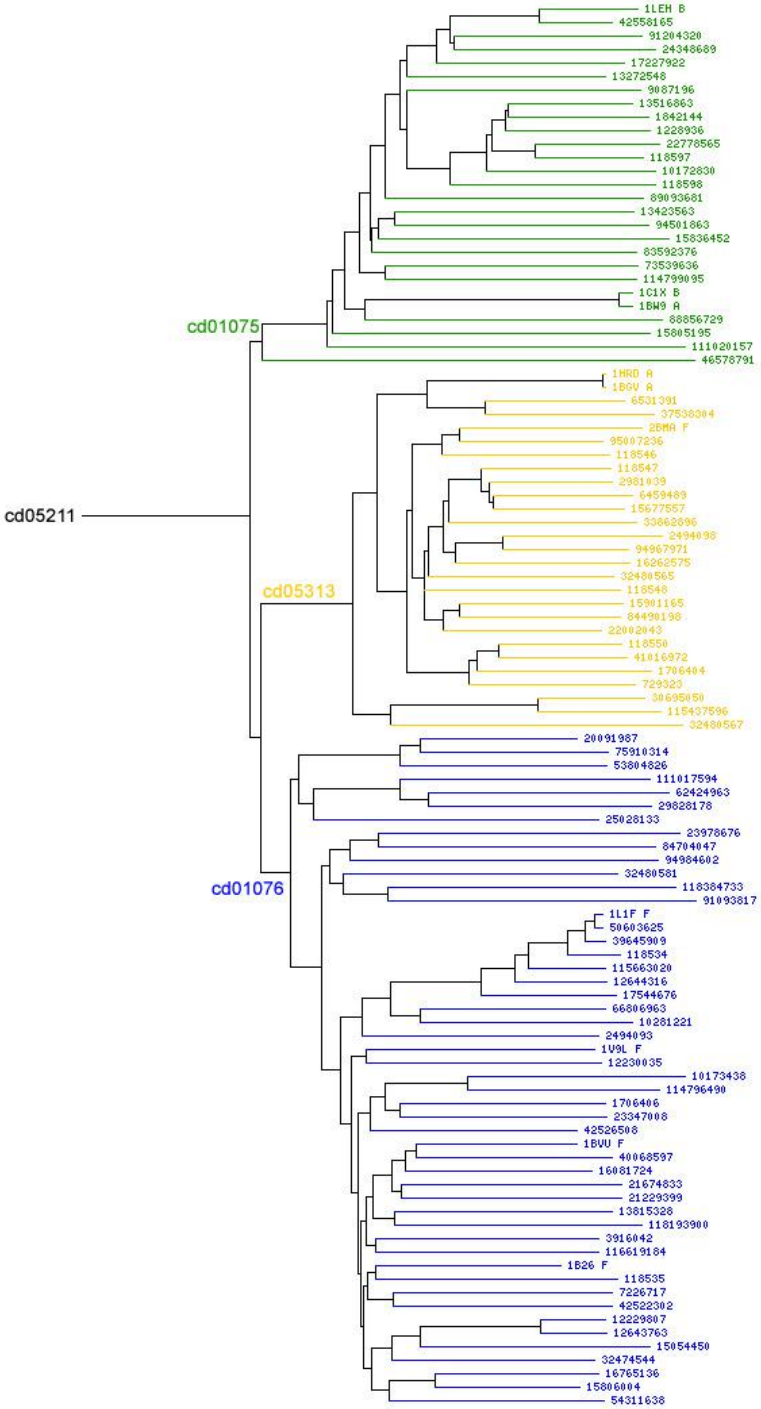


Figure A.1. Glu-Leu-Phe-Val sub-family phylogenic tree, Accession number: cd05211.

Appendix A.2 ValDH optimized codon sequence

Table A.1. *S. cinnamonensis* ValDH optimized codon sequence

1	CATATGACC	GAAGCGGAT	AACGGCGTG	CTGCATACC	CTGTTTCAT	AGCGATCAG
	GTATACTGG	CTTCGCCTA	TTGCCGCAC	GACGTATGG	GACAAAGTA	TCGCTAGTC
55	GGTGGTCAT	GAACAGGTG	GTGCTGTGC	CAGGATCGT	GCGAGCGGT	CTGAAAGCG
	CCACCAGTA	CTTGTCCAC	CACGACACG	GTCCTAGCA	CGCTCGCCA	GACTTTCGC
109	GTGATTGCG	ATTCATAGC	ACCGCGCTG	GGTCCGGCA	CTGGGTGGT	ACCCGTTTT
	CACTAACGC	TAAGTATCG	TGGCGCGAC	CCAGGCCGT	GACCCACCA	TGGGCAAAA
163	TATCCGTAT	GCGACCGAA	GAAGAAGCG	GTGGCTGAT	GTGCTGAAC	CTGAGCCGT
	ATAGGCATA	CGCTGGCTT	CTTCTTCGC	CACCGACTA	CACGACTTG	GACTCGGCA
217	GGCATGAGC	TATAAAAAC	GCGATGGCG	GGTCTGGAT	CATGGTGGG	GGTAAAGCG
	CCGTACTCG	ATATTTTTG	CGCTACCGC	CCAGACCTA	GTACCACCC	CCATTTCGC
271	GTGATTATT	GGCGATCCG	GAACAGATT	AAAAGCGAA	GATCTGCTG	CTGGCGTTT
	CACTAATAA	CCGCTAGGC	CTTGTCTAA	TTTTCGCTT	CTAGACGAC	GACCGCAAA
325	GGCCGTTTT	GTGGCGAGC	CTGGGTGGC	CGTTATGTG	ACCGCGTGC	GATGTGGGC
	CCGGCAAAA	CACCGCTCG	GACCCACCG	GCAATACAC	TGGCGCACG	CTACACCCG
379	ACCTATGTG	GCGGATATG	GATGTGGTG	GCACGTGAA	TGCCGTTGG	ACCACCGGT
	TGGATACAC	CGCCTATAC	CTACACCAC	CGTGCACTT	ACGGCAACC	TGGTGGCCA
433	CGTAGCCCG	GAAAACGGT	GGTGCGGGT	GATAGCAGC	GTGCTGACC	GCGTTTGGC
	GCATCGGGC	CTTTTGCCA	CCACGCCCA	CTATCGTCG	CACGACTGG	CGCAAACCG
487	GTGTTTCAG	GGCATGCGT	GCGAGCGCG	GAACATCTG	TGGGGCGAT	CCGAGCCTG
	CACAAAGTC	CCGTACGCA	CGCTCGCGC	CTTGTAGAC	ACCCCGCTA	GGCTCGGAC
541	CGTGGCCGT	AAAGTGGGT	GTGGCGGGT	GTGGGCAAA	GTGGGCCAT	CATCTGGTG
	GCACCGGCA	TTTCACCCA	CACCGCCCA	CACCCGTTT	CACCCGGTA	GTAGACCAC
595	GAACATCTG	CTGGAAGAT	GGTGCGGAT	GTGGTGATT	ACCGATGTG	CGTGAAGAA
	CTTGTAGAC	GACCTTCTA	CCACGCCTA	CACCACTAA	TGGCTACAC	GCACTTCTT
649	AGCGTGAAC	CGTAGCACC	CATAAACAT	CCGAGCGTG	ACCGCTGTG	GCGGATACC
	TCGCACTTG	GCATCGTGG	GTATTTGTA	GGCTCGCAC	TGGCGACAC	CGCCTATGG
703	GAAGCGCTG	ATTCGTACC	GAAGGCCTG	GATATTTAT	GCTCCGTGC	GCGCTGGGT
	CTTCGCGAC	TAAGCATGG	CTTCCGGAC	CTATAAATA	CGAGGCACG	CGCGACCCA
757	GGTGCGCTG	GATGATGAT	AGCGTGCCG	GTGCTGACC	GCTAAAGTG	GTGTGCGGT
	CCACGCGAC	CTACTACTA	TCGCACGGC	CACGACTGG	CGATTTCAC	CACACGCCA
811	GCGGCAAAC	AACCAGCTG	GCGCATCCG	GGTGTGGAA	AAAGATCTG	GCGGATCGT
	CGCCGTTTG	TTGGTCGAC	CGCGTAGGC	CCACACCTT	TTTCTAGAC	CGCCTAGCA
865	AGCATTCTG	TATGCTCCG	GATTATGTG	GTGAACGCA	GGTGGCGTG	ATTCAGGTG
	TCGTAAGAC	ATACGAGGC	CTAATACAC	CACTTGCGT	CCACCGCAC	TAAGTCCAC
919	GCGGATGAA	CTGCGTGGC	TTTGATTTT	GATCGTTGC	AAAGCGAAA	GCGAGCAAA
	CGCCTACTT	GACGCACCG	AAACTAAAA	CTAGCAACG	TTTCGCTTT	CGCTCGTTT
973	ATTTTTGAT	ACCACCCTG	GCGATTTTT	GCGCGTGCG	AAAGAAGAT	GGCATTCCT
	TAAAAACTA	TGGTGGGAC	CGCTAAAAA	CGCGCACGC	TTTCTTCTA	CCGTAAGGA
1027	CCGGCTGCG	GCTGCGGAT	CGTATTGCG	GAACAGCGT	ATGAGCGAT	GCGCGTTAA
	GGCCGACGC	CGACGCCTA	GCATAACGC	CTTGTCGCA	TACTCGCTA	CGCGCAATT
1081	AAGCTT					
	TTCGAA					

Appendix A.3 Example output of SS-BOND in *Rhodococcus sp.* M4, PDB: 1C1D

THE FOLLOWING RESIDUE PAIRS MIGHT FORM A DISULFIDE
BRIDGE ACCORDING TO THEIR CB-CB DISTANCE. WHICH
LIES BETWEEN 3.15 AND 4.50 ANGSTROM
CORRESPONDING TO CHI-3 ANGLES OF +/- 90, +/- 65.5 DEGREES

NR	RES1	--	RES2	NAME1	NAME2	CB DIST	CA DIST
1	5	--	50	ALA	ALA	4.454	5.915
2	5	--	53	ALA	LEU	3.961	6.224
3	10	--	30	GLY	ASP	4.314	4.301
4	12	--	29	MET	LEU	4.446	5.162
5	14	--	27	VAL	ILE	3.989	4.922
6	16	--	25	ARG	PHE	4.269	5.065
7	22	--	84	GLY	PRO	3.442	5.439
8	24	--	82	HIS	ALA	3.885	5.026
9	26	--	56	VAL	ALA	4.493	7.057
10	26	--	80	VAL	VAL	4.384	5.083
11	28	--	56	ARG	ALA	4.139	5.979
12	28	--	60	ARG	ALA	4.155	6.223
13	29	--	37	LEU	ALA	4.275	6.067
14	29	--	77	LEU	GLY	4.431	4.211
15	30	--	60	ASP	ALA	3.856	5.409
16	30	--	76	ASP	GLY	4.479	5.522
17	31	--	64	SER	THR	4.408	6.310
18	32	--	75	THR	GLY	4.076	4.389
19	35	--	74	GLY	MET	4.301	4.082
20	37	--	75	ALA	GLY	4.030	3.908
21	37	--	77	ALA	GLY	3.193	4.927
22	38	--	74	ALA	MET	4.353	6.422
23	38	--	114	ALA	TRP	4.489	4.724
24	39	--	63	GLY	MET	4.107	6.319
25	39	--	77	GLY	GLY	4.310	4.861
26	39	--	78	GLY	LYS	3.870	4.516
27	40	--	116	GLY	GLY	4.294	4.118
28	40	--	118	GLY	ASP	3.178	5.051
29	41	--	79	THR	SER	4.151	4.515
30	42	--	118	ARG	ASP	4.153	4.880
31	43	--	81	ALA	ILE	4.143	4.587
32	43	--	120	ALA	ASN	3.860	5.390
33	46	--	51	TYR	ASP	4.315	6.247
34	46	--	52	TYR	ALA	4.364	6.125
35	52	--	82	ALA	ALA	4.161	6.391

Appendix A.4 Sequence alignment of LeuDH, PheDHs and ValDH

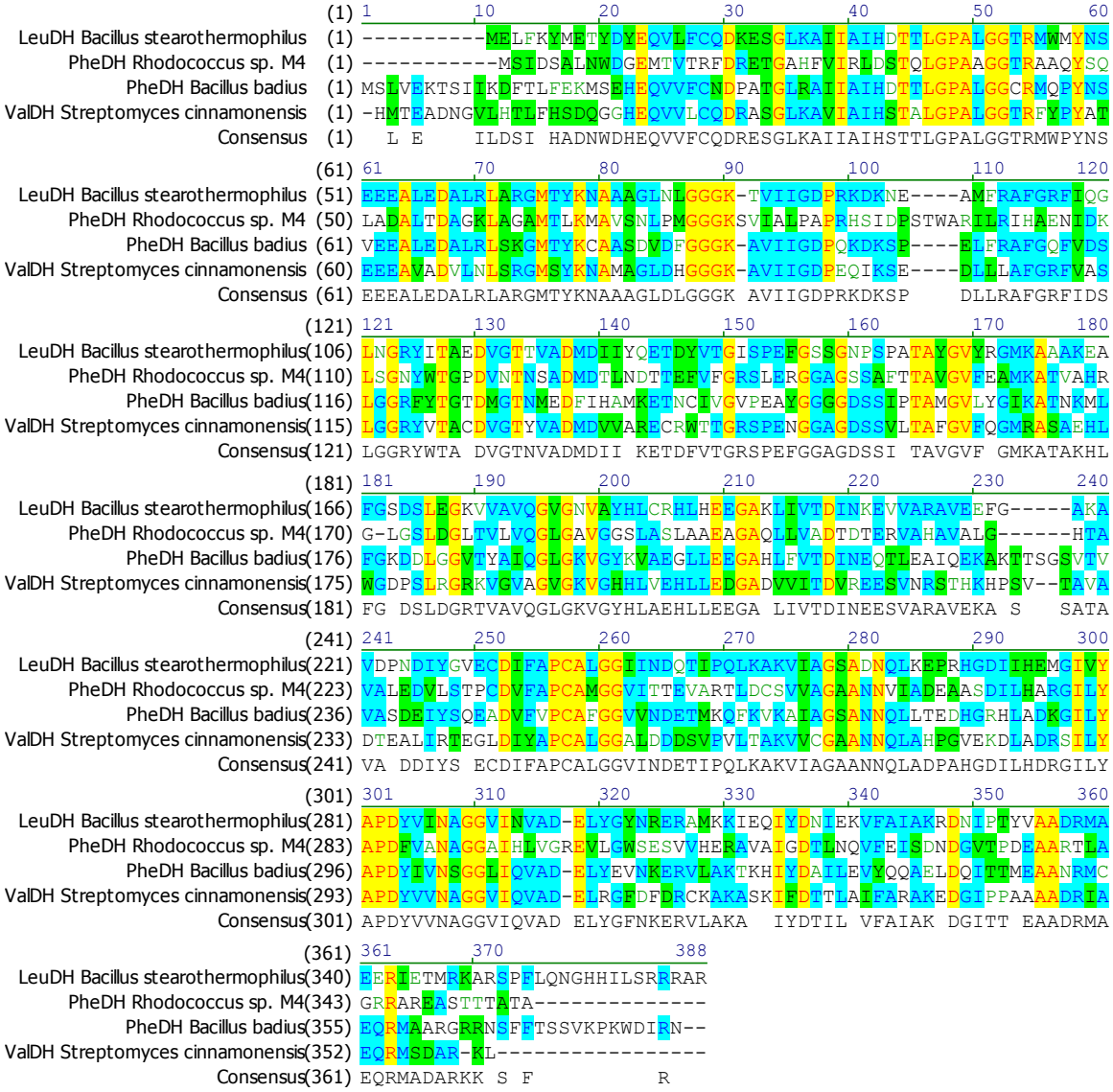


Figure A.2. A sequence alignment of *B. stearothermophilus* LeuDH, *Rh. sp.* M4 PheDH, *B. badius* PheDH, and *S. cinnamonensis* ValDH.

REFERENCES

1. WOLFENDEN, R. and SNIDER, M.J. 2001, "The depth of chemical time and the power of enzymes as catalysts." Accounts of Chemical Research, 34, (12): 938-945.

2. THAYER, A.M. 2007, "Centering on chirality." Chemical and Engineering News, 85, (32): 11-19.

3. POLLARD, D.J. and WOODLEY, J.M. 2007, "Biocatalysis for pharmaceutical intermediates: The future is now." Trends in Biotechnology, 25, (2): 66-73.

4. CAREY, J.S., LAFFAN, D., THOMSON, C., and WILLIAMS, M.T. 2006, "Analysis of the reactions used for the preparation of drug candidate molecules." Organic & Biomolecular Chemistry, 4, (12): 2337-2347.

5. HEINZE, B., KOURIST, R., FRANSSON, L., HULT, K., and BORNSCHEUER, U.T. 2007, "Highly enantioselective kinetic resolution of two tertiary alcohols using mutants of an esterase from *Bacillus subtilis*." Protein Engineering Design and Selection, 20, (3): 125-131.

6. REETZ, M.T. and SCHIMOSSEK, K. 1996, "Lipase-catalyzed dynamic kinetic resolution of chiral amines: Use of palladium as the racemization catalyst." CHIMIA International Journal for Chemistry, 50, (12): 668-669.

7. SCHMIDT, M., HASENPUSCH, D., KÄHLER, M., KIRCHNER, U., WIGGENHORN, K., LANGEL, W., and BORNSCHEUER, U.T. 2006, "Directed evolution of an esterase from *Pseudomonas fluorescens* yields a mutant with excellent enantioselectivity and activity for the kinetic resolution of a chiral building block." Chembiochem, 7, (5): 805-809.

8. BOMMARIUS, A.S. and RIEBEL, B.R., 2004."Biocatalysis." Wiley-VCH Verlag GmbH & Co.

9. CONSTABLE, D.J.C., DUNN, P.J., HAYLER, J.D., HUMPHREY, G.R., LEAZER, J.L., LINDERMAN, R.J., LORENZ, K., MANLEY, J., PEARLMAN, B.A., WELLS, A., ZAKS, A., and ZHANG, T.Y. 2007, "Key green chemistry research areas - a perspective from pharmaceutical manufacturers." Green Chemistry, 9, (5): 411-420.

10. FOX, R.J., DAVIS, S.C., MUNDORFF, E.C., NEWMAN, L.M., GAVRILOVIC, V., MA, S.K., CHUNG, L.M., CHING, C., TAM, S., MULEY, S., GRATE, J., GRUBER, J., WHITMAN, J.C., SHELDON, R.A., and HUISMAN, G.W. 2007, "Improving catalytic function by ProSAR-driven enzyme evolution." Nature Biotechnology, 25, (3): 338-344.

11. MEYER, M.M., HOCHREIN, L., and ARNOLD, F.H. 2006, "Structure-guided schema recombination of distantly related beta-lactamases." Protein Engineering Design & Selection, 19, (12): 563-570.

12. SIEGEL, J.B., ZANGHELLINI, A., LOVICK, H.M., KISS, G., LAMBERT, A.R., CLAIR, J.L.S., GALLAHER, J.L., HILVERT, D., GELB, M.H., STODDARD, B.L., HOUK, K.N., MICHAEL, F.E., and BAKER, D. 2010, "Computational design of an enzyme catalyst for a stereoselective bimolecular diels-alder reaction." Science, 329, (5989): 309-313.

13. VÁZQUEZ-FIGUEROA , E., CHAPARRO-RIGGERS , J., and BOMMARIUS, A.S. 2007, "Development of a thermostable glucose dehydrogenase by a structure-guided consensus concept." Chembiochem, 8, (18): 2295-2301.

14. LIANG, J., LALONDE, J., BORUP, B., MITCHELL, V., MUNDORFF, E., TRINH, N., KOCHREKAR, D.A., NAIR CHERAT, R., and PAI, G.G. 2009, "Development of a biocatalytic process as an alternative to the alpha-DIP-CL-mediated asymmetric reduction of a key intermediate of montelukast." Organic Process Research & Development, 14, (1): 193-198.

15. LIU, J., HSU, C.C., and WONG, C.H. 2004, "Sequential aldol condensation catalyzed by DERA mutant Ser238Asp and a formal total synthesis of atorvastatin." Tetrahedron Letters, 45, (11): 2439-2441.

16. MÜLLER, M. 2005, "Chemoenzymatic synthesis of building blocks for statin side chains." Angewandte Chemie International Edition, 44, (3): 362-365.

17. SAVILE, C.K., JANEY, J.M., MUNDORFF, E.C., MOORE, J.C., TAM, S., JARVIS, W.R., COLBECK, J.C., KREBBER, A., FLEITZ, F.J., BRANDS, J., DEVINE, P.N., HUISMAN, G.W., and HUGHES, G.J. 2010, "Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture." Science, 329, (5989): 305-309.

18. SHAFIEE, A., MOTAMEDI, H., and KING, A. 1998, "Purification, characterization and immobilization of an NADPH-dependent enzyme involved in the chiral specific reduction of the keto ester m, an intermediate in the synthesis of an anti-asthma drug, montelukast, from *Microbacterium campoquemadoensis* (MB5614)." Applied Microbiology and Biotechnology, 49, (6): 709-717.

19. MORLEY, K.L. and KAZLAUSKAS, R.J. 2005, "Improving enzyme properties: When are closer mutations better?" Trends Biotechnol, 23, (5): 231-7.

20. AHARONI, A., AMITAI, G., BERNATH, K., MAGDASSI, S., and TAWFIK, D.S. 2005, "High-throughput screening of enzyme libraries: Thiolactonases evolved by fluorescence-activated sorting of single cells in emulsion compartments." Chem Biol, 12, (12): 1281-9.

21. JEZ, J.M., BOWMAN, M.E., DIXON, R.A., and NOEL, J.P. 2000, "Structure and mechanism of the evolutionarily unique plant enzyme chalcone isomerase." Nat Struct Mol Biol, 7, (9): 786-791.

22. JEZ, J.M. and NOEL, J.P. 2002, "Reaction mechanism of chalcone isomerase." Journal of Biological Chemistry, 277, (2): 1361-1369.

23. BOMMARIUS, A.S., BLUM, J.K., and ABRAHAMSON, M.J. 2010, "Status of protein engineering for biocatalysts: How to design an industrially useful biocatalyst." Current Opinion in Chemical Biology, 15, (2): 194-200.

24. LIANG, J., MUNDORFF, E., VOLADRI, R., JENNE, S., GILSON, L., CONWAY, A., KREBBER, A., WONG, J., HUISMAN, G., TRUESDELL, S., and LALONDE, J. 2009, "Highly enantioselective reduction of a small heterocyclic ketone: Biocatalytic reduction of tetrahydrothiophene-3-one to the corresponding (R)-alcohol." Organic Process Research & Development, 14, (1): 188-192.

25. VAZQUEZ-FIGUEROA, E., YEH, V., BROERING, J.M., CHAPARRO-RIGGERS, J.F., and BOMMARIUS, A.S. 2008, "Thermostable variants constructed via the structure-guided consensus method also show increased stability in salts solutions and homogeneous aqueous-organic media." Protein Eng Des Sel, 21, (11): 673-80.

26. VAZQUEZ-FIGUEROA, E., CHAPARRO-RIGGERS, J., and BOMMARIUS, A.S. 2007, "Development of a thermostable glucose dehydrogenase by a structure-guided consensus concept." Chembiochem, 8, (18): 2295-2301.

27. MA, S.K., GRUBER, J., DAVIS, C., NEWMAN, L., GRAY, D., WANG, A., GRATE, J., HUISMAN, G.W., and SHELDON, R.A. 2010, "A green-by-design biocatalytic process for atorvastatin intermediate." Green Chemistry, 12, (1): 81-86.

28. ROTHLISBERGER, D., KHERSONSKY, O., WOLLACOTT, A.M., JIANG, L., DECHANCIE, J., BETKER, J., GALLAHER, J.L., ALTHOFF, E.A., ZANGHELLINI, A., DYM, O., ALBECK, S., HOUK, K.N., TAWFIK, D.S., and BAKER, D. 2008, "Kemp elimination catalysts by computational enzyme design." Nature, 453, (7192): 190-U4.

29. LIU, L., MURPHY, P., BAKER, D., and LUTZ, S. 2010, "Computational design of orthogonal nucleoside kinases." Chemical Communications, 46, (46): 8803-8805.

30. HU, S., HUANG, J., MEI, L., YU, Q., YAO, S., and JIN, Z. 2010, "Altering the regioselectivity of cytochrome P450 BM-3 by saturation mutagenesis for the biosynthesis of indirubin." Journal of Molecular Catalysis B: Enzymatic, 67, (12): 29-35.

31. REETZ, M.T., WANG, L.W., and BOCOLA, M. 2006, "Directed evolution of enantioselective enzymes: Iterative cycles of casting for probing protein-sequence space." Angewandte Chemie-International Edition, 45, (8): 1236-1241.

32. HEINZELMAN, P., SNOW, C.D., SMITH, M.A., YU, X.L., KANNAN, A., BOULWARE, K., VILLALOBOS, A., GOVINDARAJAN, S., MINSHULL, J., and ARNOLD, F.H. 2009, "Schema recombination of a fungal cellulase uncovers a single mutation that contributes markedly to stability." Journal of Biological Chemistry, 284, (39): 26229-26233.

33. HEINZELMAN, P., SNOW, C.D., WU, I., NGUYEN, C., VILLALOBOS, A., GOVINDARAJAN, S., MINSHULL, J., and ARNOLD, F.H. 2009, "A family of thermostable fungal cellulases created by structure-guided recombination." Proc Natl Acad Sci U S A, 106, (14): 5610-5.

34. HEINZELMAN, P., KOMOR, R., KANAAN, A., ROMERO, P., YU, X., MOHLER, S., SNOW, C., and ARNOLD, F. 2010, "Efficient screening of fungal cellobiohydrolase class I enzymes for thermostabilizing sequence blocks by SCHEMA structure-guided recombination." Protein Eng Des Sel, 23, (11): 871-80.

35. REETZ, M.T., D CARBALLEIRA, J., and VOGEL, A. 2006, "Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability." Angewandte Chemie-International Edition, 45, (46): 7745-7751.

36. BLOOM, J.D. and ARNOLD, F.H. 2009, "In the light of directed evolution: Pathways of adaptive protein evolution." Proc Natl Acad Sci, 106: 9995-10000.

37. BOUGIOUKOU, D.J., KILLE, S., TAGLIEBER, A., and REETZ, M.T. 2009, "Directed evolution of an enantioselective enoate-reductase: Testing the utility of iterative saturation mutagenesis." Advanced Synthesis & Catalysis, 351, (18): 3287-3305.

38. REETZ, M.T. and CARBALLEIRA, J.D. 2007, "Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes." Nature Protocols, 2, (4): 891-903.

39. KOURIST, R., JOCHENS, H., BARTSCH, S., KUIPERS, R., PADHI, S.K., GALL, M., BOTTCHER, D., JOOSTEN, H.J., and BORNSCHEUER, U.T. 2010, "The alpha/beta-hydrolase fold 3DM database (ABHDB) as a tool for protein engineering." Chembiochem, 11, (12): 1635-1643.

40. KUIPERS, R.K., JOOSTEN, H.J., VAN BERKEL, W.J.H., LEFERINK, N.G.H., ROOIJEN, E., ITTMANN, E., VAN ZIMMEREN, F., JOCHENS, H., BORNSCHEUER, U., VRIEND, G., DOS SANTOS, V.A.P.M., and SCHAAP, P.J. 2010, "3DM: Systematic analysis of heterogeneous superfamily data to discover protein functionalities." Proteins-Structure Function and Bioinformatics, 78, (9): 2101-2113.

41. REETZ, M.T., KAHAKEAW, D., and LOHMER, R. 2008, "Addressing the numbers problem in directed evolution." Chembiochem, 9, (11): 1797-1804.

42. REETZ, M.T. and WU, S. 2008, "Greatly reduced amino acid alphabets in directed evolution: Making the right choice for saturation mutagenesis at homologous enzyme positions." Chemical Communications, (43): 5499-5501.

43. MENA, M.A. and DAUGHERTY, P.S. 2005, "Automated design of degenerate codon libraries." Protein Eng Des Sel, 18, (12): 559-61.

44. JOCHENS, H. and BORNSCHEUER, U.T. 2010, "Natural diversity to guide focused directed evolution." Chembiochem, 11, (13): 1861-6.

45. MEYER, M.M., SILBERG, J.J., VOIGT, C.A., ENDELMAN, J.B., MAYO, S.L., WANG, Z.G., and ARNOLD, F.H. 2003, "Library analysis of SCHEMA-guided protein recombination." Protein Science, 12, (8): 1686-1693.

46. KHERSONSKY, O., ROTHLISBERGER, D., DYM, O., ALBECK, S., JACKSON, C.J., BAKER, D., and TAWFIK, D.S. 2010, "Evolutionary optimization of computationally designed enzymes: Kemp eliminases of the KE07 series." Journal of Molecular Biology, 396, (4): 1025-1042.

47. KOSZELEWSKI, D., TAUBER, K., FABER, K., and KROUTIL, W. 2010, "Omega-transaminases for the synthesis of non-racemic alpha-chiral primary amines." Trends in Biotechnology, 28, (6): 324-332.

48. KHERSONSKY, O., ROTHLISBERGER, D., WOLLACOTT, A.M., MURPHY, P., DYM, O., ALBECK, S., KISS, G., HOUK, K.N., BAKER, D., and TAWFIK, D.S. 2011, "Optimization of the in-silico-designed kemp eliminase KE70 by computational design and directed evolution." Journal of Molecular Biology, 407, (3): 391-412.

49. PEISAJOVICH, S.G. and TAWFIK, D.S. 2007, "Protein engineers turned evolutionists." Nature Methods, 4, (12): 991-994.

50. GERLT, J.A. and BABBITT, P.C. 2009, "Enzyme (re)design: Lessons from natural evolution and computation." Current Opinion in Chemical Biology, 13, (1): 10-18.

51. KAZLAUSKAS, R.J. and BORNSCHEUER, U.T. 2009, "Finding better protein engineering strategies." Nature Chemical Biology, 5, (8): 526-529.

52. BOMMARIUS, A.S. 2010, "Protein engineering: Check nature first, then evolve." Nat Chem Biol, 6, (11): 793-794.

53. HÖHNE, M., SCHATZLE, S., JOCHENS, H., ROBINS, K., and BORNSCHEUER, U.T. 2010, "Rational assignment of key motifs for function guides *in silico* enzyme identification." Nat Chem Biol, 6, (11): 807-813.

54. REETZ, M.T., SONI, P., FERNANDEZ, L., GUMULYA, Y., and CARBALLEIRA, J.D. 2010, "Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-fit method." Chemical Communications, 46, (45): 8657-8658.

55. KIM, M.S. and LEI, X.G. 2008, "Enhancing thermostability of *Escherichia coli* phytase APPA2 by error-prone PCR." Applied Microbiology and Biotechnology, 79, (1): 69-75.

56. SILBERG, J.J., ENDELMAN, J.B., and ARNOLD, F.H. 2004, "SCHEMA-guided protein recombination." Protein Engineering, 388: 35-42.

57. BERSHTEIN, S., GOLDIN, K., and TAWFIK, D.S. 2008, "Intense neutral drifts yield robust and evolvable consensus proteins." Journal of Molecular Biology, 379, (5): 1029-1044.

58. BERSHTEIN, S. and TAWFIK, D.S. 2008, "Advances in laboratory evolution of enzymes." Current Opinion in Chemical Biology, 12, (2): 151-158.

59. HUISMAN, G.W., LIANG, J., and KREBBER, A. 2010, "Practical chiral alcohol manufacture using ketoreductases." Current Opinion in Chemical Biology, 14, (2): 122-129.

60. ALEXANDROVA, A.N., ROTHLISBERGER, D., BAKER, D., and JORGENSEN, W.L. 2008, "Catalytic mechanism and performance of computationally designed enzymes for kemp elimination." Journal of the American Chemical Society, 130, (47): 15907-15915.

61. BREUER, M., DITRICH, K., HABICHER, T., HAUER, B., KEBELER, M., STÜRMER, R., and ZELINSKI, T. 2004, "Industrial methods for the production of optically active intermediates." Angewandte Chemie International Edition, 43, (7): 788-824.

62. ITOH, N., YACHI, C., and KUDOME, T. 2000, "Determining a novel NAD+-dependent amine dehydrogenase with a broad substrate range from *Streptomyces virginiae* IFO 12827: Purification and characterization." Journal of Molecular Catalysis B: Enzymatic, 10, (13): 281-290.

63. HÖHNE, M. and BORNSCHEUER, U.T. 2009, "Biocatalytic routes to optically active amines." ChemCatChem, 1, (1): 42-51.

64. NUGENT, T.C. and EL-SHAZLY, M. 2010, "Chiral amine synthesis – recent developments and trends for enamide reduction, reductive amination, and imine reduction." Advanced Synthesis & Catalysis, 352, (5): 753-819.

65. WOODLEY, J.M. 2008, "New opportunities for biocatalysis: Making pharmaceutical processes greener." Trends in Biotechnology, 26, (6): 321-327.

66. JACQUES, J., COLLET, A., and WILEN, S.H., 1981."Enantiomers, racemates, and resolutions. New York, NY: Wiley.

67. JOHANSSON, A. 1995, "Methods for the asymmetric preparation of amines." Contemporary Organic Synthesis, 2, (6): 393-407.

68. STEWART, J.D. 2001, "Dehydrogenases and transaminases in asymmetric synthesis." Current Opinion in Chemical Biology, 5, (2): 120-129.

69. TAYLOR, P.P., PANTALEONE, D.P., SENKPEIL, R.F., and FOTHERINGHAM, I.G. 1998, "Novel biosynthetic approaches to the production of unnatural amino acids using transaminases." Trends in Biotechnology, 16, (10): 412-418.

70. TRUPPO, M.D., TURNER, N.J., and ROZZELL, J.D. 2009, "Efficient kinetic resolution of racemic amines using a transaminase in combination with an amino acid oxidase." Chem Comm, 2009: 2127-2129.

71. TUFVESSON, P., LIMA-RAMOS, J., JENSEN, J.S., AL-HAQUE, N., NETO, W., and WOODLEY, J.M. 2011, "Process considerations for the asymmetric synthesis of chiral amines using transaminases." Biotechnology and Bioengineering, 108, (7): 1479-1493.

72. KOSZELEWSKI, D., LAVANDERA, I., CLAY, D., ROZZELL, D., and KROUTIL, W. 2008, "Asymmetric synthesis of optically pure pharmacologically relevant amines employing omega-transaminases." Advanced Synthesis & Catalysis, 350, (17): 2761-2766.

73. PELLISSIER, H. 2011, "Recent developments in dynamic kinetic resolution." Tetrahedron, 67, (21): 3769-3802.

74. KOSZELEWSKI, D., LAVANDERA, I., CLAY, D., GUEBITZ, G.M., ROZZELL, D., and KROUTIL, W. 2008, "Formal asymmetric biocatalytic reductive amination." Angewandte Chemie International Edition, 47, (48): 9337-9340.

75. SHIN, J.S. and KIM, B.G. 1999, "Asymmetric synthesis of chiral amines with ω-transaminase." Biotechnology and Bioengineering, 65, (2): 206-211.

76. TRUPPO, M.D., ROZZELL, J.D., and TURNER, N.J. 2009, "Efficient production of enantiomerically pure chiral amines at concentrations of 50 g/l using transaminases." Organic Process Research & Development, 14, (1): 234-237.

77. POPOV, V.O. and LAMZIN, V.S. 1994, "NAD(+)-dependent formate dehydrogenase." The Biochemical journal, 301 (Pt 3): 625-643.

78. VAN DER DONK, W.A. and ZHAO, H. 2003, "Recent developments in pyridine nucleotide regeneration." Current Opinion in Biotechnology, 14, (4): 421-426.

79. ABRAHAMSON, M.J., VÁZQUEZ-FIGUEROA, E., WOODALL, N.B., MOORE, J.C., and BOMMARIUS, A.S. 2012, "Development of an amine dehydrogenase for synthesis of chiral amines." Angewandte Chemie International Edition, 51, (16): 3969-3972.

80. MEHLING, A., WEHMEIER, U.F., and PIEPERSBERG, W. 1995, "Application of random amplified polymorphic DNA (RAPD) assays in identifying conserved regions of actinomycete genomes." FEMS Microbiology Letters, 128, (2): 119-125.

81. DAWSON, R.M.C., ELLIOTT, D.C., and ELLIOTT, W.H., 1989."Data for biochemical research." Clarendon Press.

82. STRATAGENE, "Quikchange® site-directed mutagenesis kit". 1998.

83. SAMBROOK, J. and RUSSELL, D.W., 2001."Molecular cloning: A laboratory manual." Cold Spring Harbor Laboratory Press.

84. KAPLAN, N.O., COLOWICK, S.P., and BARNES, C.C. 1951, "Effect of alkali on diphosphopyridine nucleotide." Journal of Biological Chemistry, 191, (2): 461-472.

85. LOWRY, O.H., ROBERTS, N.R., and KAPPHAHN, J.I. 1957, "The fluorometric measurement of pyridine nucleotides." J. Biol. Chem., 224: 1047-1064.

86. TSOTSOU, G.E., CASS, A.E.G., and GILARDI, G. 2002, "High throughput assay for cytochrome P450 BM3 for screening libraries of substrates and combinatorial mutants." Biosensors and Bioelectronics, 17, (12): 119-131.

87. SEIDEMANN, J., 1973."A flexible system of enzymatic analysis." New York, NY. Academic Press.

88. CHEN, S.H. and ENGEL, P.C. 2009, "Efficient screening for new amino acid dehydrogenase activity: Directed evolution of *Bacillus sphaericus* phenylalanine dehydrogenase towards activity with an unsaturated non-natural amino acid." Journal of Biotechnology, 142, (2): 127-134.

89. BAKER, P.J., TURNBULL, A.P., SEDELNIKOVA, S.E., STILLMAN, T.J., and RICE, D.W. 1995, "A role for quaternary structure in the substrate specificity of leucine dehydrogenase." Structure, 3, (7): 693-705.

90. BRUNHUBER, N.M.W., THODEN, J.B., BLANCHARD, J.S., and VANHOOKE, J.L. 2000, "*Rhodococcus* L-phenylalanine dehydrogenase: Kinetics, mechanism, and structural basis for catalytic specifity." Biochemistry, 39, (31): 9174-9187.

91. VANHOOKE, J.L., THODEN, J.B., BRUNHUBER, N.M.W., BLANCHARD, J.S., and HOLDEN, H.M. 1999, "Phenylalanine dehydrogenase from *Rhodococcus sp.* M4: High-resolution x-ray analyses of inhibitory ternary complexes reveal key features in the oxidative deamination mechanism." Biochemistry, 38, (8): 2326-2339.

92. SEKIMOTO, T., MATSUYAMA, T., FUKUI, T., and TANIZAWA, K. 1993, "Evidence for lysine 80 as general base catalyst of leucine dehydrogenase." Journal of Biological Chemistry, 268, (36): 27039-27045.

93. VÁZQUEZ-FIGUEROA, E., "Development of a novel dehydrogenase and a stable cofactor regeneration system." 2008, Georgia Institute of Technology: Atlanta, GA.

94. REETZ, M.T., BOCOLA, M., CARBALLEIRA, J.D., ZHA, D.X., and VOGEL, A. 2005, "Expanding the range of substrate acceptance of enzymes: Combinatorial active-site saturation test." Angewandte Chemie-International Edition, 44, (27): 4192-4196.

95. CHEN, S. and ENGEL, P.C. 2007, "An engineered mutant, L307V of phenylalanine dehydrogenase from *Bacillus sphaericus*: High activity and stability in organic-aqueous solvent mixtures and utility for synthesis of non-natural L-amino acids." Enzyme and Microbial Technology, 40, (5): 1407-1411.

96. SUPELCO, "Perfluoro acid anhydrides". 1997, Sigma-Aldrich: Bellefonte, PA.

97. LANDON, S.R., CHAN, Y.M., SONOLA, O.O., and TATCHELL, A.R. 1984, "Asymmetric syntheses. Part 11. Reduction of ketones and related ketone oximes with lithium aluminum hydride-3-O-cyclohexylmethyl-1,2-O-cyclohexylidene-α-D-glucofuranose complex to give optically active alcohols and amines." Journal of the Chemical Society, Perkin Transactions 1, (3): 493-6.

98. NAGATA, S., TANIZAWA, K., ESAKI, N., SAKAMOTO, Y., OHSHIMA, T., TANAKA, H., and SODA, K. 1988, "Gene cloning and sequence determination of leucine dehydrogenase from *Bacillus stearothermophilus* and structural comparison with other NAD(P)+-dependent dehydrogenases." Biochemistry, 27, (25): 9056-9062.

99. AGRANAT, I. and WAINSCHTEIN, S.R. 2010, "The strategy of enantiomer patents of drugs." Drug Discovery Today, 15, (56): 163-170.

100. BRUNHUBER, N.M.W. and BLANCHARD, J.S. 1994, "The biochemistry and enzymology of amino-acid dehydrogenases." Critical Reviews in Biochemistry and Molecular Biology, 29, (6): 415-467.

101. KATAOKA, K., TAKADA, H., TANIZAWA, K., YOSHIMURA, T., ESAKI, N., OHSHIMA, T., and SODA, K. 1994, "Construction and characterization of chimeric enzyme consisting of an amino-terminal domain of phenylalanine dehydrogenase and a carboxy-terminal domain of leucine dehydrogenase." Journal of Biochemistry, 116, (4): 931-936.

102. LIU, Z.-J., SUN, Y.-J., ROSE, J., CHUNG, Y.-J., HSIAO, C.-D., CHANG, W.-R., KUO, I., PEROZICH, J., LINDAHL, R., HEMPEL, J., and WANG, B.-C. 1997, "The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD and the Rossmann fold." Nat Struct Mol Biol, 4, (4): 317-326.

103. MARCHLER-BAUER, A., LU, S., ANDERSON, J.B., CHITSAZ, F., DERBYSHIRE, M.K., DEWEESE-SCOTT, C., FONG, J.H., GEER, L.Y., GEER, R.C., GONZALES, N.R., GWADZ, M., HURWITZ, D.I., JACKSON, J.D., KE, Z., LANCZYCKI, C.J., LU, F., MARCHLER, G.H., MULLOKANDOV, M., OMELCHENKO, M.V., ROBERTSON, C.L., SONG, J.S., THANKI, N., YAMASHITA, R.A., ZHANG, D., ZHANG, N., ZHENG, C., and BRYANT, S.H. 2010, "CDD: A conserved domain database for the functional annotation of proteins." Nucleic Acids Research.

104. ASANO, Y., NAKAZAWA, A., and ENDO, K. 1987, "Novel phenylalanine dehydrogenases from *Sporosarcina ureae* and *Bacillus sphaericus*. Purification and characterization." Journal of Biological Chemistry, 262, (21): 10346-54.

105. ASANO, Y., NAKAZAWA, A., ENDO, K., HIBINO, Y., OHMORI, M., NUMAO, N., and KONDO, K. 1987, "Phenylalanine dehydrogenase of *Bacillus badius*." European Journal of Biochemistry, 168, (1): 153-159.

106. COOPER, A.J.L., LEUNG, L.K.H., and ASANO, Y. 1989, "Enzymatic cycling assay for phenylpyruvate." Analytical Biochemistry, 183, (2): 210-214.

107. HUMMEL, W., SCHÜTTE, H., SCHMIDT, E., WANDREY, C., and KULA, M.R. 1987, "Isolation of L-phenylalanine dehydrogenase from *Rhodococcus sp.* M4 and its application for the production of L-phenylalanine." Applied Microbiology and Biotechnology, 26, (5): 409-416.

108. SEAH, S.Y.K., LINDA BRITTON, K., BAKER, P.J., RICE, D.W., ASANO, Y., and ENGEL, P.C. 1995, "Alteration in relative activities of phenylalanine dehydrogenase towards different substrates by site-directed mutagenesis." FEBS Letters, 370, (12): 93-96.

109. VILLALONGA, R., FUJII, A., SHINOHARA, H., TACHIBANA, S., and ASANO, Y. 2008, "Covalent immobilization of phenylalanine dehydrogenase on cellulose membrane for biosensor construction." Sensors and Actuators B: Chemical, 129, (1): 195-199.

110. VILLALONGA, R., TACHIBANA, S., CAO, R., MATOS, M., and ASANO, Y. 2007, "Glycosidation of phenylalanine dehydrogenase with O-carboxymethyl-poly-beta-cyclodextrin." Enzyme and Microbial Technology, 40, (3): 471-475.

111. VILLALONGA, R., TACHIBANA, S., PÉREZ, Y., and ASANO, Y. 2005, "Increased conformational and thermal stability properties for phenylalanine dehydrogenase by chemical glycosidation with end-group activated dextran." Biotechnology Letters, 27, (17): 1311-1317.

112. MAITI, R., VAN DOMSELAAR, G.H., ZHANG, H., and WISHART, D.S. 2004, "Superpose: A simple server for sophisticated structural superposition." Nucleic Acids Research, 32, (Web Server Issue): 590-594.

113. UNITED STATES DEPARTMENT OF JUSTICE and DRUG ENFORCEMENT ADMINISTRATION. "Controlled substances list". 2012, 7/24/2012; http://www.deadiversion.usdoj.gov/schedules/orangebook/c_cs_alpha.pdf.

114. JAENICKE, R. 1991, "Protein stability and molecular adaptation to extreme conditons." European Journal of Biochemistry, 202, (3): 715-728.

115. MATTHEWS, B.W., NICHOLSON, H., and BECKTEL, W.J. 1987, "Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding." Proceedings of the National Academy of Sciences, 84, (19): 6663-6667.

116. CORNISH-BOWDEN, A., 2004."Fundamentals of enzyme kinetics." Portland Press.

117. PAVIA, D.L., 2009."Introduction to spectroscopy." Brooks/Cole, Cengage Learning.

118. ARMSTRONG, D.W., RUNDLETT, K.L., and NAIR, U.B. 1996, "Enantioresolution of amphetamine, methamphetamine, and deprenyl (selegiline) by LC, GC and CE." Current Separations, 15, (2): 57-61.

119. PLENEVAUX, A., DEWEY, S.L., FOWLER, J.S., GUILLAUME, M., and WOLF, A.P. 1990, "Synthesis of (R)-(-)- and (S)-(+)-4-fluorodeprenyl, (R)-(-)- and (S)-(+)-[N-11C-methyl]-4-fluorodeprenyl and pet studies in baboon brain." Journal of Medicinal Chemistry, 33, (7): 2015-2019.

120. NAVARRETE, R.M., VARA, J.S.A., and HUTCHINSON, C.R. 1990, "Purification of an inducible L-valine dehydrogenase of *Streptomyces coelicolor* A3(2)." Journal of General Microbiology, 136, (2): 273-281.

121. TANG, L. and HUTCHINSON, C.R. 1993, "Sequence, transcriptional, and functional analyses of the valine (branched-chain amino acid) dehydrogenase gene of *Streptomyces coelicolor*." Journal of Bacteriology, 175, (13): 4176-4185.

122. TURNBULL, A.P., BAKER, P.J., and RICE, D.W. 1997, "Analysis of the quaternary structure, substrate specificity, and catalytic mechanism of valine dehydrogenase." Journal of Biological Chemistry, 272, (40): 25105-25111.

123. NGUYEN, L.T., NGUYEN, K.T., SPIZEK, J., and BEHAL, V. 1995, "The tylosin producer, *Streptomyces fradiae*, contains a second valine dehydrogenase." Microbiology, 141, (5): 1139-1145.

124. TANG, L., ZHANG, Y.X., and HUTCHINSON, C.R. 1994, "Amino acid catabolism and antibiotic synthesis: Valine is a source of precursors for macrolide biosynthesis in *Streptomyces ambofaciens* and *Streptomyces fradiae*." Journal of Bacteriology, 176, (19): 6107-6119.

125. VANCURA, A., VANCUROVA, I., VOLC, J., FUSSEY, S., FLIEGER, M., NEUZIL, J., MARSAKEL, J., and BEHAL, V. 1988, "Valine dehydrogenase from *Streptomyces fradiae*: Purification and properties." Journal of General Microbiology, 134, (12): 3213-3219.

126. LEISER, A., BIRCH, A., and ROBINSON, J.A. 1996, "Cloning, sequencing, overexpression in *Escherichia coli*, and inactivation of the valine dehydrogenase gene in the polyether antibiotic producer *Streptomyces cinnamonensis*." Gene, 177, (12): 217-222.

127. PRIESTLEY, N.D. and ROBINSON, J.A. 1989, "Purification and catalytic properties of L-valine dehydrogenase from *Streptomyces cinnamonensis*." The Biochemical journal, 261, (3): 853-861.

128. ROBINSON, M., LILLEY, R., LITTLE, S., EMTAGE, J.S., YARRANTON, G., STEPHENS, P., MILLICAN, A., EATON, M., and HUMPHREYS, G. 1984, "Codon usage can affect efficiency of translation of genes in *Escherichia coli*." Nucleic Acids Research, 12, (17): 6663-6671.

129. WADA, K., WADA, Y., ISHIBASHI, F., GOJOBORI, T., and IKEMURA, T. 1992, "Codon usage tabulated from the Genbank genetic sequence data." Nucleic Acids Research, 20: 2111-2118.

130. HAMMES-SCHIFFER, S. and BENKOVIC, S.J. 2006, "Relating protein motion to catalysis." Annual Review of Biochemistry, 75, (1): 519-541.

131. WONG, K.F., WATNEY, J.B., and HAMMES-SCHIFFER, S. 2004, "Analysis of electrostatics and correlated motions for hydride transfer in dihydrofolate reductase." The Journal of Physical Chemistry B, 108, (32): 12231-12241.

132. MANSFELD, J., VRIEND, G., DIJKSTRA, B.W., VELTMAN, O.R., VAN DEN BURG, B., VENEMA, G., ULBRICH-HOFMANN, R., and EIJSINK, V.G.H. 1997, "Extreme stabilization of a thermolysin-like protease by an engineered disulfide bond." Journal of Biological Chemistry, 272, (17): 11152-11156.

133. FILIKOV, A.V., HAYES, R.J., LUO, P., STARK, D.M., CHAN, C., KUNDU, A., and DAHIYAT, B.I. 2002, "Computational stabilization of human growth hormone." Protein Science, 11, (6): 1452-1461.

134. KORKEGIAN, A., BLACK, M.E., BAKER, D., and STODDARD, B.L. 2005, "Computational thermostabilization of an enzyme." Science, 308, (5723): 857-860.

135. SHEFFLER, W. and BAKER, D. 2009, "RosettaHoles: Rapid assessment of protein core packing for structure prediction, refinement, design, and validation." Protein Science, 18, (1): 229-239.

136. SHEFFLER, W. and BAKER, D. 2010, "RosettaHoles2: A volumetric packing measure for protein structure refinement and validation." Protein Science, 19, (10): 1991-1995.

137. STRICKLER, S.S., GRIBENKO, A.V., GRIBENKO, A.V., KEIFFER, T.R., TOMLINSON, J., REIHLE, T., LOLADZE, V.V., and MAKHATADZE, G.I. 2006, "Protein stability and surface electrostatics: A charged relationship." Biochemistry, 45, (9): 2761-2766.

138. WUNDERLICH, M., MARTIN, A., and SCHMID, F.X. 2005, "Stabilization of the cold shock protein CSPB from *Bacillus subtilis* by evolutionary optimization of coulombic interactions." Journal of Molecular Biology, 347, (5): 1063-1076.

139. HAZES, B. and DIJKSTRA, B.W. 1988, "Model building of disulfide bonds in proteins with known three-dimensional structure." Protein Engineering, 2, (2): 119-125.

140. O'CONNOR, B.D. and YEATES, T.O. 2004, "GDAP: A web tool for genome-wide protein disulfide bond prediction." Nucleic Acids Research, 32, (Web Server Issue): W360-W364.

141. BORGO, B. and HAVRANEK, J.J. 2012, "Automated selection of stabilizing mutations in designed and natural proteins." Proceedings of the National Academy of Sciences, 109, (5): 1494-1499.

142. SLUSARCZYK, H., FELBER, S., KULA, M.-R., and POHL, M. 2000, "Stabilization of NAD-dependent formate dehydrogenase from *Candida boidinii* by site-directed mutagenesis of cysteine residues." European Journal of Biochemistry, 267, (5): 1280-1289.

143. FABER, K., 1992."Biotransformations in organic chemistry." Springer-Verlag.

144. MOHAMMAD, A., 2012."Green solvents I: Properties and applications in chemistry." Springer.

145. AMERICAN CHEMICAL SOCIETY, G.C.I., PHARMACEUTICAL ROUNDTABLE. "ACS GCI pharmaceutical roundtable solvent selection guide". 2011 April 1, 2011; www.acs.org.gcipharmaroundtable.

146. CAINELLI, G., ENGEL, P.C., GALLETTI, P., GIACOMINI, D., GUALANDI, A., and PARADISI, F. 2005, "Engineered phenylalanine dehydrogenase in organic solvents: Homogeneous and biphasic enzymatic reactions." Organic & Biomolecular Chemistry, 3, (24): 4316-4320.

147. BOMMARIUS, A.S., "Reduction of C=N bonds," Enzyme catalysis in organic synthesis. 2002, Wiley-VCH. p. 1047-1063.

148. MENZEL, A., WERNER, H., ALTENBUCHNER, J., and GROGER, H. "From enzymes to "Designer bugs" In reductive amination: A new process for the synthesis of l-tert-leucine using a whole cell-catalyst." 2004.

149. WILMS, B., HAUCK, A., REUSS, M., SYLDATK, C., MATTES, R., SIEMANN, M., and ALTENBUCHNER, J. 2001, "High-cell-density fermentation for production of *n*-carbamoylase using an expression system based on the *Escherichia coli* RHABAD promoter." Biotechnology and Bioengineering, 73, (2): 95-103.

150. WOLFGANG BERKE, HANS-JÜRGEN SCHÜZ, CHRISTIAN WANDREY, MICHAEL MORR, GUDRUN DENDA, and MARIA-REGINA KULA 1988, "Continuous regeneration of ATP in enzyme membrane reactor for enzymatic syntheses." Biotechnology and Bioengineering, 32, (2): 130-139.

VITA

MICHAEL J. ABRAHAMSON

Michael Abrahamson was born in Saint Petersburg, Florida, and attended public school in Seminole, Florida. He graduated *summa cum laude* from the University of Florida, Gainesville, Florida in 2007 with a B.A. in Chemical Engineering before coming to Georgia Tech to pursue a doctorate in Chemical Engineering.